

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 18:25:06 ; Search time 50 Seconds
(without alignments)
1812.657 Million cell updates/sec

Title: US-09-715-927-6
Perfect score: 2929
Sequence: 1 MTRAGDHNRQCCGLADY.....ACGPAKEVRKENQANTSVW 571

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :			
A_Geneseq_13Jun03.*			
1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*		
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*		
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*		
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*		
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*		
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*		
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*		
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*		
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*		
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*		
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*		
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*		
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*		
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*		
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*		
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*		
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*		
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*		
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*		
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*		
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*		
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*		
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*		
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	2929	100.0	571 22 ABUS2922	Cell structure and
2	2929	100.0	571 22 AAM39134	Human polypeptide
3	2929	100.0	571 22 AAB27245	Human EXMAD-23 SEQ
4	2925	99.9	571 23 AAE23866	Human ferroportin
5	2917	99.6	571 20 AAY29327	Human secreted pro
6	2917	99.6	571 22 AAU39053	Human secreted pro
7	2917	99.6	571 23 ABB55762	Human polypeptide
8	2018	68.9	408 23 ABP41875	Human ovarian anti
9	1789	61.1	382 23 ABB89115	Human polypeptide

10	1789	61.1	383	19	AAW74869	Human secreted pro
11	1789	61.1	383	23	ABG95320	Human novel secret
12	1630	55.7	439	22	AAM40920	Human polypeptide
13	1611	55.0	439	22	ABG20262	Novel human diagno
14	1470.5	50.2	319	22	ABUS2922	Cell structure and
15	859	29.3	167	22	ABUS2922	Human protein sequ
16	606	20.7	123	21	ABG93563	Human secreted pro
17	302	10.3	576	23	ABG93606	Herbicideally activ
18	295	10.1	150	24	ABG75821	Transposers and i
19	181.5	6.2	371	21	ABG09603	Arabidopsis thalia
20	159	5.4	242	21	ABG09604	Arabidopsis thalia
21	157	5.4	496	23	ABP65787	Bifidobacterium lo
22	143	4.9	231	21	ABG09605	Arabidopsis thalia
23	128	4.4	1313	22	ABG63904	Drosophila melanog
24	123.5	4.2	524	23	ABG76953	Human protein, hom
25	123.5	4.2	524	23	ABB05602	Human glucose tran
26	120	4.1	421	23	ABP66017	Bifidobacterium lo
27	120	4.1	458	22	AAU03207	Fruit fly G protei
28	119	4.1	537	22	AAG92976	C glutamicum prote
29	115.5	3.9	491	22	ABW70144	Drosophila melanog
30	115.5	3.9	683	22	AAU44669	Propionibacterium
31	115	3.9	538	22	AAG83020	S. epidermidis ope
32	114.5	3.9	424	22	AAG89887	C glutamicum prote
33	114.5	3.9	424	22	AAB78954	C. glutamicum SKT
34	114.5	3.9	424	22	AAB76809	Corynebacterium gl
35	114	3.9	392	21	AAI81756	Streptococcus pneu
36	114	3.9	392	24	ABU01535	S. pneumoniae type
37	113.5	3.9	373	19	AAW80675	S. pneumoniae prot
38	113.5	3.9	522	18	AAW17836	Rat glucose transp
39	113	3.9	475	22	ABBS7889	Drosophila melanog
40	113	3.9	522	22	AAB86970	D. melanogaster pe
41	112	3.8	475	22	ABB67155	Drosophila melanog
42	111.5	3.8	512	23	AAO14190	Human transporter
43	109.5	3.7	470	23	ABP39765	Staphylococcus epi
44	109.5	3.7	480	21	AAG28044	Arabidopsis thalia
45	109.5	3.7	483	21	AAG28043	Arabidopsis thalia

ALIGNMENTS

RESULT 1
ABUS2922
ID ABUS2922 standard; Protein; 571 AA.
XX
AC ABUS2922;
XX
DT 15-APR-2003 (first entry)
XX
DE Cell structure and motility-associated protein from DKFZphutel_24j6.
XX
KW Human; gene therapy; vaccine; disease treatment; detection.
XX
OS Homo sapiens.
XX
PN WO200112659-A2.
XX
PD 22-FEB-2001.
XX
PF 18-AUG-2000; 2000WO-IB01496.
XX
PR 18-AUG-1999; 99US-0149499.
PR 28-SEP-1999; 99US-0156503.
XX
(GEHU-) GERMAN HUMAN GENOME PROJECT.
XX
PA Wiemann S;
XX
PI WPI; 2001-327840/34.
DR N-PSDB; ABX71330.
XX
PT Nucleic acids having the sequences of clones isolated from libraries of different human tissues, useful in recombinant DNA methodologies -

XX Claim 21; Page 537; 1095pp; English.

XX This invention describes novel polynucleotides and polypeptides isolated from human cDNA libraries which can be used for gene therapy or in vaccines. The polynucleotides of the invention and antibodies encoded by them may be used in the prevention, diagnosis and treatment of diseases associated with inappropriate polypeptide expression. The products of the invention may also be used to identify modulators of expression and activity and to down regulate expression and activity. The antibodies of the invention may also be used as diagnostic agents for detecting the presence of polypeptides in samples. This sequence represents a polypeptide described in the disclosure of the invention.

XX Sequence 571 AA;

Query Match 100.0%; Score 2929; DB 22; Length 571;
Best Local Similarity 100.0%; Pred. No. 8.7e-271;
Matches 571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTRAGDHNQRQCGCGSLADYLTSAKFLYLGHSLSTWGDWMHFAVSFLVELYGNLSLL 60
Db 1 MTRAGDHNQRQCGCGSLADYLTSAKFLYLGHSLSTWGDWMHFAVSFLVELYGNLSLL 60

Qy 61 TAVYGLVWAGSVLVGLGAIIGDWDKNARLKVAQTSLVVQNVSVILCGIILMMVFLKHHEL 120
Db 61 TAVYGLVWAGSVLVGLGAIIGDWDKNARLKVAQTSLVVQNVSVILCGIILMMVFLKHHEL 120

Qy 121 LTMVHGWLTSYLLIITIANIANLASTATAITIQDWMIVVVGEDRSKLANNNATIRRI 180
Db 121 LTMVHGWLTSYLLIITIANIANLASTATAITIQDWMIVVVGEDRSKLANNNATIRRI 180

Qy 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGNLVSMCVVEYVLLWKVYQKTPALAVKAGLK 240
Db 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGNLVSMCVVEYVLLWKVYQKTPALAVKAGLK 240

Qy 241 EEETELQNLNKHKTEPKLEGTHLMGVKDSNIHELEHEQEPPTCASQMAEPFRFDGWV 300
Db 241 EEETELQNLNKHKTEPKLEGTHLMGVKDSNIHELEHEQEPPTCASQMAEPFRFDGWV 300

Qy 301 SYNQPVFLAGMGLAFYMTVLGFDCTTGAYYTGQISGLSILMGASAITGIMGTVAF 360
Db 301 SYNQPVFLAGMGLAFYMTVLGFDCTTGAYYTGQISGLSILMGASAITGIMGTVAF 360

Qy 361 TWLRKGLVRTGLISGLAQSLILCVISVFMGSPDLDSVPEDIRSRFTQGESITP 420
Db 361 TWLRKGLVRTGLISGLAQSLILCVISVFMGSPDLDSVPEDIRSRFTQGESITP 420

Qy 421 TKIPEITTEIYMSNGSANSIVPETSPEVPIISVSLFAGVIAARIGLWSFDLTVTQLL 480
Db 421 TKIPEITTEIYMSNGSANSIVPETSPEVPIISVSLFAGVIAARIGLWSFDLTVTQLL 480

Qy 481 QENVIESERGIINGVQNSMNYLLDLHFIWVILAPNPEAFGLLVLSVSFVAMGHIMYFR 540
Db 481 QENVIESERGIINGVQNSMNYLLDLHFIWVILAPNPEAFGLLVLSVSFVAMGHIMYFR 540

Qy 541 FAQNTLGNKLPACGPDPAKEVRKENQANTSV 571
Db 541 FAQNTLGNKLPACGPDPAKEVRKENQANTSV 571

RESULT 2
AAM39134
ID AAM39134 standard; Protein; 571 AA.
XX AAM39134;
AC AAM39134;
XX AAM39134;
DT 22-OCT-2001 (first entry)
XX Human polypeptide SEQ ID NO 2279.
XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW

Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
anyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukaemia.
Homo sapiens.
WO200153312-A1.
26-JUL-2001.
26-DEC-2000; 2000WO-US34263.
21-JAN-2000; 2000US-0488725.
25-APR-2000; 2000US-0552317.
09-JUL-2000; 2000US-0598042.
19-JUL-2000; 2000US-0620312.
03-AUG-2000; 2000US-0653450.
14-SEP-2000; 2000US-0662191.
19-OCT-2000; 2000US-0693036.
29-NOV-2000; 2000US-0727344.
(HYSE-) HYSEQ INC.
Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
Zhao QA, Zhou P, Goodrich R, Drmanac RT;
WPI; 2001-442253/47.
DR N-PSDB; AAI58290.
Novel nucleic acids and polypeptides, useful for treating disorders
such as central nervous system injuries -
Example 4; SEQ ID NO 2279; 10078pp; English.
The invention relates to human nucleic acids (AAI57798-AAI61369) and
the encoded polypeptides (AAM38642-AAM42213) with nontropic,
immunosuppressant and cytostatic activity. The polynucleotides are useful
in gene therapy. A composition containing a polypeptide or polynucleotide
of the invention may be used to treat diseases of the peripheral nervous
system, such as peripheral nervous injuries, peripheral neuropathy and
centralised neuropathies and central nervous system diseases, such as
Alzheimer's, Parkinson's disease, Huntington's disease, anyotrophic
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
and thrombolytic activity, cancer diagnosis and therapy, drug screening,
assays for receptor activity, arthritis and inflammation, leukaemias and
C.N.S disorders.
Note: The sequence data for this patent did not form part of the printed
specification.

Query Match 100.0%; Score 2929; DB 22; Length 571;
Best Local Similarity 100.0%; Pred. No. 8.7e-271;
Matches 571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTRAGDHNQRQCGCGSLADYLTSAKFLYLGHSLSTWGDWMHFAVSFLVELYGNLSLL 60
Db 1 MTRAGDHNQRQCGCGSLADYLTSAKFLYLGHSLSTWGDWMHFAVSFLVELYGNLSLL 60

Qy 61 TAVYGLVWAGSVLVGLGAIIGDWDKNARLKVAQTSLVVQNVSVILCGIILMMVFLKHHEL 120
Db 61 TAVYGLVWAGSVLVGLGAIIGDWDKNARLKVAQTSLVVQNVSVILCGIILMMVFLKHHEL 120

Qy 121 LTMVHGWLTSYLLIITIANIANLASTATAITIQDWMIVVVGEDRSKLANNNATIRRI 180
Db 121 LTMVHGWLTSYLLIITIANIANLASTATAITIQDWMIVVVGEDRSKLANNNATIRRI 180

Qy 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGNLVSMCVVEYVLLWKVYQKTPALAVKAGLK 240
Db 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGNLVSMCVVEYVLLWKVYQKTPALAVKAGLK 240

Db 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGMNLSMCMVEYVLLWKVYOKTPALAVKAGLK 240
QY 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPRTFRDGMV 300
Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPRTFRDGMV 300
QY 301 SYNOPVFLAGMGLAFYMTVLGFCITTGVAYTQGLSGSILSLMGASAITGIMGTVA 360
Db 301 SYNOPVFLAGMGLAFYMTVLGFCITTGVAYTQGLSGSILSLMGASAITGIMGTVA 360
QY 361 TWLRKCGLVRTGLISGLAQSLCLILCVISVPMFGSPDLVSPPEDIRSRFIQGESITP 420
Db 361 TWLRKCGLVRTGLISGLAQSLCLILCVISVPMFGSPDLVSPPEDIRSRFIQGESITP 420
QY 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLLPAGVIAARIGLWGFDLTVTQLL 480
Db 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLLPAGVIAARIGLWGFDLTVTQLL 480
QY 481 QENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
Db 481 QENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
QY 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
Db 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571

RESULT 3
AAB27245
ID AAB27245 standard; Protein; 571 AA.
XX
AC AAB27245;
XX
DT 27-MAR-2001 (first entry)
XX
DE Human EXMAD-23 SEQ ID NO: 23.
XX
KW Extracellular matrix and adhesion-associated protein; EXMAD; cancer;
KW inflammation; reproductive disorder; cardiovascular disorder;
KW immune disorder; musculoskeletal disorder; developmental disorder;
KW gastrointestinal disorder; cell proliferation disorder.
XX
OS Homo sapiens.
XX
PN WO200068380-A2.
XX
PD 16-NOV-2000.
XX
PF 10-MAY-2000; 2000WO-US12811.
XX
PR 11-MAY-1999; 99US-0133643.
XX
PR 23-AUG-1999; 99US-0150409.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Bandman O, Hillman JL, Tang YT, Lai P, Yue H, Baughn MR, Lu DAM;
PI Azimzai Y;
XX
DR WPI; 2001-007395/01.
DR N-PSDB; AAC66912.
XX
XX Isolated polynucleotide encoding extracellular matrix or
PT adhesion-associated protein (EXMAD) useful for diagnosing, treating, or
PT preventing disorders associated with expression of EXMAD such as
PT proliferative, immune and genetic disorders -
XX
PS Claim 1; Page 109-110; 129pp; English.
XX
CC The present invention provides the protein and coding sequences for 25
CC novel extracellular matrix and adhesion-associated proteins (EXMADS).
CC These are designated EXMAD-1, EXMAD-2, EXMAD-3, EXMAD-4, EXMAD-5,
CC EXMAD-6, EXMAD-7, EXMAD-8, EXMAD-9, EXMAD-10, EXMAD-11, EXMAD-12,
CC EXMAD-13, EXMAD-14, EXMAD-15, EXMAD-16, EXMAD-17, EXMAD-18, EXMAD-19,

CC EXMAD-20, EXMAD-21, EXMAD-22, EXMAD-23, EXMAD-24 and EXMAD-25. They are
CC useful in the prevention and treatment of cancers, cell proliferation,
CC cardiovascular, reproductive, immune, musculoskeletal, developmental and
CC gastrointestinal disorders and inflammation.
XX
SQ Sequence 571 AA;
Query Match 100.0%; Score 2929; DB 22; Length 571;
Best Local Similarity 100.0%; Pred. No. 8.7e-271;
Matches 571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTRAGDHNRRQGCCGSLADYLTSAKFLLYLGHSLSTWGDMMHFAVSFLVELYNSLLL 60
Db 1 MTRAGDHNRRQGCCGSLADYLTSAKFLLYLGHSLSTWGDMMHFAVSFLVELYNSLLL 60
QY 61 TAVYGLVAGSVLVGLGIDVDKNARLKAQVQSLVVQNVSVILCGIILMMVFLKHHEL 120
Db 61 TAVYGLVAGSVLVGLGIDVDKNARLKAQVQSLVVQNVSVILCGIILMMVFLKHHEL 120
QY 121 LTMVHGWLVTSCYILITIANIANLASTATAITIQRDWIVVVGEDRSKLANMNAITIRRI 180
Db 121 LTMVHGWLVTSCYILITIANIANLASTATAITIQRDWIVVVGEDRSKLANMNAITIRRI 180
QY 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGMNLSMCMVEYVLLWKVYOKTPALAVKAGLK 240
Db 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGMNLSMCMVEYVLLWKVYOKTPALAVKAGLK 240
QY 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPRTFRDGMV 300
Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPRTFRDGMV 300
QY 301 SYNOPVFLAGMGLAFYMTVLGFCITTGVAYTQGLSGSILSLMGASAITGIMGTVA 360
Db 301 SYNOPVFLAGMGLAFYMTVLGFCITTGVAYTQGLSGSILSLMGASAITGIMGTVA 360
QY 361 TWLRKCGLVRTGLISGLAQSLCLILCVISVPMFGSPDLVSPPEDIRSRFIQGESITP 420
Db 361 TWLRKCGLVRTGLISGLAQSLCLILCVISVPMFGSPDLVSPPEDIRSRFIQGESITP 420
QY 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLLPAGVIAARIGLWGFDLTVTQLL 480
Db 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLLPAGVIAARIGLWGFDLTVTQLL 480
QY 481 QENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
Db 481 QENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
QY 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
Db 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
RESULT 4
AAB23866
ID AAB23866 standard; Protein; 571 AA.
XX
AC AAB23866;
XX
DT 10-SEP-2002 (first entry)
XX
DE Human ferroportin 1 mutant protein.
XX
KW Human; ferroportin 1; hereditary haemochromatosis; therapy; mutant;
KW mutein.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Misc-difference 77
FT /note= "Wild type Ala substituted with Asp"
XX
PN WO200233119-A2.

XX PD 25-APR-2002.
 XX PF 17-OCT-2001; 2001WO-EPI2018.
 XX PR 17-OCT-2000; 2000IT-MI02240.
 XX PA (PIET/) PIETRANGELO A.
 XX PI Pietrangelo A;
 XX XX WPI; 2002-444246/47.
 DR DR N-PSDB; AAD38291.
 XX XX Novel nucleic acid coding for ferroportin 1 useful for the in vitro
 PT diagnosis of hereditary hemochromatosis in a mammal, comprises a
 PT mutation of the codon coding for the ferroportin 1
 XX
 PS Claim 13; Page 35-36; 37pp; English.
 XX
 CC The invention relates to mutations in the gene coding for human
 CC ferroportin 1 associated with hereditary hemochromatosis and methods
 CC for the diagnosis of hereditary hemochromatosis based on the
 CC identification of such mutations. Polynucleotides of the invention
 CC are used for the production of recombinant mutated ferroportin 1.
 CC its fragment or a chimeric protein including the fragment, in order
 CC to study the functional characteristics of the mutated protein. The present
 CC are also used for diagnostic and therapeutic purposes. The present
 CC sequence is human ferroportin 1 mutant protein.
 XX
 SQ Sequence 571 AA;
 Query Match 99.9%; Score 2925; DB 23; Length 571;
 Best Local Similarity 99.8%; Pred. No. 2.1e-270;
 Matches 570; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MTRAGDHNRGCGSLADYLTSAKPLLYLGHSLSTWGDHMHFAVSFLVLYGNSLL 60
 DB 1 MTRAGDHNRGCGSLADYLTSAKPLLYLGHSLSTWGDHMHFAVSFLVLYGNSLL 60
 QY 61 TAVYGLVAGSVLVGLGIIIGDWDKNAKLVQTSLVVQNVSVILCGIILMVFLLKHKL 120
 DB 61 TAVYGLVAGSVLVGLGIIIGDWDKNAKLVQTSLVVQNVSVILCGIILMVFLLKHKL 120
 QY 121 LTMVHGVLNLSVLIITIANIANLASTATAITIQDWIIVVAGEDRSKLANNNATIRI 180
 DB 121 LTMVHGVLNLSVLIITIANIANLASTATAITIQDWIIVVAGEDRSKLANNNATIRI 180
 QY 181 DQLTNIIAPWAVGQIMTFGSPVIGCGFISGNLVSMCVSVLLWKVYKTPALAVKAGLK 240
 DB 181 DQLTNIIAPWAVGQIMTFGSPVIGCGFISGNLVSMCVSVLLWKVYKTPALAVKAGLK 240
 QY 241 EEETELKQLNLHKDTEPKLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPPTFRDGV 300
 DB 241 EEETELKQLNLHKDTEPKLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPPTFRDGV 300
 QY 301 SYYNQPVFLAGMLAFIYMTVGFDCITTCYAYTQGLSGSILILMGASAITGIMGTAF 360
 DB 301 SYYNQPVFLAGMLAFIYMTVGFDCITTCYAYTQGLSGSILILMGASAITGIMGTAF 360
 QY 361 TWLRKCGLVRTGLISGLAQLSCILICVISVFMGSPDLDSVSPFEDIRSRFTQGESITP 420
 DB 361 TWLRKCGLVRTGLISGLAQLSCILICVISVFMGSPDLDSVSPFEDIRSRFTQGESITP 420
 QY 421 TKIPEITTEIYMGNSNSANIVPETSPEPIISVSLFAGVIAARIGLSFDTVTQLL 480
 DB 421 TKIPEITTEIYMGNSNSANIVPETSPEPIISVSLFAGVIAARIGLSFDTVTQLL 480
 QY 481 QENVIESERGIINGVQNSMYLLDLHLHFIMVILAPNPEAFGLLVLSVSVFAMGHIMYFR 540
 DB 481 QENVIESERGIINGVQNSMYLLDLHLHFIMVILAPNPEAFGLLVLSVSVFAMGHIMYFR 540
 QY 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571

Db 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
 RESULT 5
 AAY29327
 ID AAY29327 standard; Protein; 571 AA.
 XX AC AAY29327;
 XX DT 29-SEP-1999 (first entry)
 XX DE Human secreted protein clone ewl50_1 protein sequence.
 XX KW Human; secreted protein; nutrition; cytokine; cytokine; cell proliferation;
 KW differentiation; immune stimulating; vaccine; suppression; gene therapy;
 KW haematopoiesis regulation; tissue growth; activin; inhibin; cadherin;
 KW chemotactic; chemokinetic; haemostatic; thrombolytic; anti-inflammatory;
 KW tumour invasion suppressor; tumour inhibition.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT Misc-difference 202 /label= unknown
 FT /note= "encoded by RTC"
 FT Misc-difference 504
 FT /label= unknown
 FT /note= "encoded by RAT"
 XX WO9937674-A1.
 XX PD 29-JUL-1999.
 XX PF 21-JAN-1999; 99WO-US01404.
 XX PR 20-JAN-1999; 99US-0235609.
 XX PR 22-JAN-1998; 98US-0072134.
 XX (GEMY) GENETICS INST INC.
 XX Agostino MJ, Clark HF, Collins-Racie LA, Fecthel K;
 PI Jacobs K, Lavallie ER, McCoy JM, Merberg D, Steininger RJ;
 PI Treacy M, Wong GG;
 XX WPI; 1999-458682/38.
 DR N-PSDB; AAX90442.
 XX New polynucleotides encoding secreted human proteins derived from,
 PT e.g. fetal brain potentially used as immunostimulators
 XX Claim 16; Page 115-117; 139pp; English.
 CC The present sequence represents a human secreted protein. Human secreted
 CC protein polynucleotides and proteins are predicted to have biological
 CC activities which would make them suitable for treating, preventing or
 CC ameliorating medical conditions in humans and animals, although no
 CC supporting data is given. Suggested activities include nutritional
 CC activity, cytokine and cell proliferation/differentiation activity,
 CC immune stimulating (e.g. as vaccines) or suppressing activity,
 CC haematopoiesis regulating activity, tissue growth activity,
 CC activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, cadherin/tumour invasion suppressor activity, and tumour
 CC inhibition activity. The polynucleotides are also stated to be useful
 CC for gene therapy.
 XX Sequence 571 AA;
 Query Match 99.6%; Score 2917; DB 20; Length 571;
 Best Local Similarity 99.6%; Pred. No. 1.2e-269;
 Matches 569; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MTRAGDHNRRQCGGSLADYLTSAKFLLYLGHSLSTWGDNRWHPFAVSFLVELYGNLSLL 60
Db 1 MTRAGDHNRRQCGGSLADYLTSAKFLLYLGHSLSTWGDNRWHPFAVSFLVELYGNLSLL 60
QY 61 TAVYGLVAVAGSVLVGLGAIIGDWDVKNARLKVAQTSILVQNVSVILCGIILMMVFLHKHEL 120
Db 61 TAVYGLVAVAGSVLVGLGAIIGDWDVKNARLKVAQTSILVQNVSVILCGIILMMVFLHKHEL 120
QY 121 LTMWYHGWLTSCYIIITITIANIANLASTATAITIQDWMIVVAGEDRSKLANMATIRRI 180
Db 121 LTMWYHGWLTSCYIIITITIANIANLASTATAITIQDWMIVVAGEDRSKLANMATIRRI 180
QY 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGWNLSVMCEVYLLWKVYQKTPALAVKAGLK 240
Db 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGWNLSVMCEVYLLWKVYQKTPALAVKAGLK 240
QY 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHEHEQEPTCASQMAEPFRTRDGMV 300
Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHEHEQEPTCASQMAEPFRTRDGMV 300
QY 301 SYYNQPVFLAGMGLAFLYMTVLGFDCTITTGAYTQGLSGSILSILMGASAITGIMTVAF 360
Db 301 SYYNQPVFLAGMGLAFLYMTVLGFDCTITTGAYTQGLSGSILSILMGASAITGIMTVAF 360
QY 361 TWLRKCKGLVRLTGLISGLAQLSCLLVCISLVFMPGSPDLVSPDPEDIRSRFIQGESITP 420
Db 361 TWLRKCKGLVRLTGLISGLAQLSCLLVCISLVFMPGSPDLVSPDPEDIRSRFIQGESITP 420
QY 421 TKIPBITTEIYMSNGSNANIPETSPESVPIISVLSLFPAGVIARIGLWSFDTLVTQLL 480
Db 421 TKIPBITTEIYMSNGSNANIPETSPESVPIISVLSLFPAGVIARIGLWSFDTLVTQLL 480
QY 481 QENVIESBERGLINGVQNSMNVLLDLHFIMVILAPNPEAFGLLVLSVSVFVANGHIMYFR 540
Db 481 QENVIESBERGLINGVQNSMNVLLDLHFIMVILAPNPEAFGLLVLSVSVFVANGHIMYFR 540
QY 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
Db 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
```

RESULT 6

AAU33053
ID AAU33053 standard; Protein; 571 AA.

AC AAU33053;

DT 16-JAN-2002 (first entry)

XX Human secreted protein ew150_1.

XX Human; secreted protein; antiinflammatory; immunosuppressive;
KW neutropic; neuroprotective; antiarthritis; antimicrobial; vulnery;
KW cystostatic; antidiabetic; viricide; antinfertility; anticonvulsant;
KW vasotropic; antiparkinsonian; immunostimulant; dermatological;
KW antirheumatic; antitumor; antulcer; osteopathic; tranquiliser;
KW cerebroprotective; cytokine; cell proliferation; cell differentiation;
KW immune deficiency; severe combined immunodeficiency; SCID; tumour;
KW autoimmune disorder; multiple sclerosis; rheumatoid arthritis;
KW graft-versus-host disease; myeloid deficiency; wound healing; ulcer;
KW periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease;
KW Parkinson's disease; Huntington's disease; infection; cardiac disease;
KW stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
KW food supplement; vaccine.

XX Homo sapiens.

XX WO200175068-A2.

XX 11-OCT-2001.

XX 22-MAR-2001; 2001WO-US09369.

XX

PR 30-MAR-2000; 2000US-0539330.
XX 04-DEC-2000; 2000US-0729674.
XX (GEMY) GENETICS INST INC.
XX Jacobs K, McCoy JM, Lavallie E, Collins-racie LA, Evans C;
PI Treacy M, Agostino MJ, Steininger RJ, Spaulding V, Wong GG;
PI Clark H, Fichtel K, Merberg D;
XX WPI; 2001-639363/73.
DR N-PSDB; AAS59271.
XX

Secreted human proteins, useful as vaccine for treating various diseases such as autoimmune disorders (e.g. multiple sclerosis), and nervous system disorders (e.g. stroke)

Disclosure; Page 542-544; 619pp; English.

The invention relates to novel human secreted proteins, the nucleic acids encoding them. The protein may exhibit cytokine, cell proliferation or cell differentiation activity or may induce production of other cytokines in certain cell populations and may exhibit immune stimulating or immune suppressing activity, which is useful for the treatment of various immune deficiencies and disorders e.g. severe combined immunodeficiency (SCID), autoimmune disorders e.g. multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation. The proteins are also useful in the treatment of diseases and disorders including tissue, skin and organ transplantation and in graft-versus-host diseases (GVHD), in the induction of tumour immunity, myeloid or lymphoid cell deficiencies, wound healing and tissue repair, in the treatment of burns, incisions and ulcers; as well as in treatment of periodontal disease, osteoporosis or osteoarthritis, mediated by inflammatory processes, diseases of the peripheral nervous system, Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome, infections, infarction of cardiac and central nervous system vessel e.g. stroke, sepsis, inflammatory bowel disease, ulcers, bone regeneration. The protein, having activin- or inhibin-related activities is useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. The proteins and nucleic acids are also useful as food supplements. The present sequence represents a secreted protein of the invention.

SQ Sequence 571 AA;

Query Match 99.6%; Score 2917; DB 22; Length 571;
Best Local Similarity 99.6%; Pred. No. 1.2e-269;
Matches 569; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTRAGDHNRRQCGGSLADYLTSAKFLLYLGHSLSTWGDNRWHPFAVSFLVELYGNLSLL 60
Db 1 MTRAGDHNRRQCGGSLADYLTSAKFLLYLGHSLSTWGDNRWHPFAVSFLVELYGNLSLL 60
QY 61 TAVYGLVAVAGSVLVGLGAIIGDWDVKNARLKVAQTSILVQNVSVILCGIILMMVFLHKHEL 120
Db 61 TAVYGLVAVAGSVLVGLGAIIGDWDVKNARLKVAQTSILVQNVSVILCGIILMMVFLHKHEL 120
QY 121 LTMWYHGWLTSCYIIITITIANIANLASTATAITIQDWMIVVAGEDRSKLANMATIRRI 180
Db 121 LTMWYHGWLTSCYIIITITIANIANLASTATAITIQDWMIVVAGEDRSKLANMATIRRI 180
QY 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGWNLSVMCEVYLLWKVYQKTPALAVKAGLK 240
Db 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGWNLSVMCEVYLLWKVYQKTPALAVKAGLK 240

QY 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHEHEQEPTCASQMAEPFRTRDGMV 300
Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHEHEQEPTCASQMAEPFRTRDGMV 300
QY 301 SYYNQPVFLAGMGLAFLYMTVLGFDCTITTGAYTQGLSGSILSILMGASAITGIMTVAF 360
Db 301 SYYNQPVFLAGMGLAFLYMTVLGFDCTITTGAYTQGLSGSILSILMGASAITGIMTVAF 360

QY 361 TWLRKCGLVRTGLISGLAQLCLILCVISVFMFGSPDLDSVSPFFDIRSRFIQGESITP 420
Db |||||
QY 361 TWLRKCGLVRTGLISGLAQLCLILCVISVFMFGSPDLDSVSPFFDIRSRFIQGESITP 420
Db |||||
QY 421 TKIPEITTEIYMGNSANSANIVPETSPEVPIISVSLLFAGVIAARIGLWSFDLTVTQLL 480
Db |||||
QY 481 QENVIESERGIINGVQNMNLLDLHLFIMVILAPNPEARGLVLVLSVSVAMGHIMYFR 540
Db |||||
QY 541 FAQNTLGNKLPACPDPAKEVRKENQANTSVV 571
Db |||||
ID ABB55762 standard; Protein; 571 AA.
AC ABB55762;
XX
DT 14-FEB-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 130.
XX
KW Human; clone bd306-7; clone yb8-1; ATCC number 98599; gene therapy;
KW immune disorder; bacterial infection; fungal infection; cancer; tumour;
KW autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibin;
KW osteoporosis; osteoarthritis; nervous system disorder; neuropathy;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;
KW haemophilia; cardiac infarction; stroke; sepsis; arthritis; vulnery;
KW ischaemia-reperfusion injury; inflammatory bowel disease; chemotactic;
KW Crohn's disease; cytostatic; anti-inflammatory; immunomodulator;
KW neuroprotective; haemostatic; thrombolytic; anti-inflammatory.
OS Homo sapiens.
XX
FN US2001039335-A1.
XX
PD 08-NOV-2001.
XX
PF 04-DEC-2000; 2000US-0729674.
XX
PR 10-APR-1997; 97US-126435P.
PR 04-DEC-1997; 97US-067454P.
PR 20-DEC-1997; 97US-068379P.
PR 02-JAN-1998; 98US-070346P.
PR 07-JAN-1998; 98US-070643P.
PR 08-JAN-1998; 98US-070755P.
PR 13-JAN-1998; 98US-071304P.
PR 22-JAN-1998; 98US-072134P.
PR 30-JAN-1998; 98US-073095P.
PR 18-FEB-1998; 98US-075038P.
PR 30-MAR-2000; 2000US-0539330.
PR 23-NOV-1998; 98US-0197886.
XX
PA (JACO/) JACOBS K.
PA (MCCO/) MCCOY J M.
PA (LAVA/) LAVALLIE E R.
PA (COLL/) COLLINS-RACIE L A.
PA (EVAN/) EVANS C.
PA (MERB/) MERBERG D.
PA (TREA/) TREACY M.
PA (AGOS/) AGOSTINO M J.
PA (STEI/) STEININGER R J.
PA (SPAU/) SPAULDING V.
PA (WONG/) WONG G G.
PA (CLAR/) CLARK H.
PA (FECH/) FECHTEL K.
XX
PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;

PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
PI Wong GG, Clark H, Fechtel K;
XX
DR WPI; 2002-040725/05.
DR N-PSDB; ABA90940.
XX
PT New secreted proteins and encoding polynucleotides, useful in gene
PT therapies, particularly for preventing or treating autoimmune
PT disorders, cancer, graft-versus-host disease, wound, osteoporosis,
PT stroke or inflammations
XX
PS Disclosure; Page 270-271; 349pp; English.
XX
CC The invention relates to isolated polynucleotides (ABA90876-ABA90968 and
CC ABA90980) and encoded proteins (ABBS5698-ABBS5800), especially
CC polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and
CC proteins SEQ ID NO 2 (ABBS5698) and SEQ ID NO 20 (ABBS5707) contained in
CC clones bd306-7 and yb8-1 respectively and the clones bd306-7 and yb8-1
CC are deposited with the American Type Culture Collection (ATCC) with
CC accession number 98599. The polynucleotides and encoded polypeptides have
CC cytostatic, anti-inflammatory, immunomodulator, vulnerary,
CC neuroprotective, activin, inhibin, chemotactic, haemostatic, thrombolytic
CC and anti-inflammatory activity and acting as cytokine modulators,
CC haematopoiesis regulators, tissue growth modulators and/or cadherin
CC suppressors. The polypeptides and polynucleotides are useful in gene
CC therapies, particularly for preventing, treating or ameliorating any of
CC the following diseases: immune deficiency and disorders; e.g. bacterial
CC or fungal infections, autoimmune disorders, cancer, systemic lupus
CC erythematosus or graft-versus-host disease; myeloid or lymphoid cell
CC deficiencies; wound, burns, incisions and ulcers, osteoporosis or
CC osteoarthritis; central and peripheral nervous system diseases and
CC neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's
CC disease, amyotrophic lateral sclerosis or Shy-brager syndrome;
CC haemophilia, cardiac infarction or stroke, inflammations, shock, sepsis
CC or systemic inflammatory response syndrome, ischaemia-reperfusion
CC injury, endotoxin lethality, arthritis, inflammatory bowel disease or
CC Crohn's disease; or tumours or cancers, pemphigus vulgaris or pemphigus
CC foliaceus.
XX
SQ Sequence 571 AA;
Query Match 99.6%; Score 2917; DB 23; Length 571;
Best Local Similarity 99.6%; Pred. No. 1.2e-269;
Matches 569; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MTRAGDHNRRQCCGSLADYLTSAKFLYLGHSLSTWGDMMHFAVSFLVLYGNSLLL 60
Db |||||
QY 1 MTRAGDHNRRQCCGSLADYLTSAKFLYLGHSLSTWGDMMHFAVSFLVLYGNSLLL 60
Db |||||
QY 61 TAVYGLVWAGSVLVLGAIGDWDKNARLKVQAOTSLVQVNSVILCGIILMMVFLHKHEL 120
Db |||||
QY 61 TAVYGLVWAGSVLVLGAIGDWDKNARLKVQAOTSLVQVNSVILCGIILMMVFLHKHEL 120
Db |||||
QY 121 LTMHGWLTSCYIIITIANIANLASTATAITIQDMMIVVVGAGSDRSKLANNNATIRI 180
Db |||||
QY 121 LTMHGWLTSCYIIITIANIANLASTATAITIQDMMIVVVGAGSDRSKLANNNATIRI 180
Db |||||
QY 181 DOLTNILAPMAVGQIMTFGSPVIGCGFISGMNLVSNKVEYVLLWKVYQKTPALAVKAGLK 240
Db |||||
QY 181 DOLTNILAPMAVGQIMTFGSPVIGCGFISGMNLVSNKVEYVLLWKVYQKTPALAVKAGLK 240
Db |||||
QY 241 BEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPRTFRDGVV 300
Db |||||
QY 241 BEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPRTFRDGVV 300
Db |||||
QY 301 SYYNQPVFLAGMGLAFLYMTVLGFCITTTGAYTQGLSGSILSLMGASAITGIMGTAF 360
Db |||||
QY 301 SYYNQPVFLAGMGLAFLYMTVLGFCITTTGAYTQGLSGSILSLMGASAITGIMGTAF 360
Db |||||
QY 361 TWLRKCGLVRTGLISGLAQLCLILCVISVFMFGSPDLDSVSPFFDIRSRFIQGESITP 420
Db |||||
QY 361 TWLRKCGLVRTGLISGLAQLCLILCVISVFMFGSPDLDSVSPFFDIRSRFIQGESITP 420
Db |||||

QY 421 TKIPETITTYMNGNSANIVPETSPEVPIISVLSLLFAGVIAARIGLWSFDLTVTQLL 480
Db 421 TKIPETITTYMNGNSANIVPETSPEVPIISVLSLLFAGVIAARIGLWSFDLTVTQLL 480
QY 481 QENVIESERGIINGVQNSMNYLLDLHLHFMVILAPNPEAFGLLVLSVSFVAMGHIMYPR 540
Db 481 QENVIESERGIINGVQNSMNYLLDLHLHFMVILAPNPEAFGLLVLSVSFVAMGHIMYPR 540
QY 541 FQAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
Db 541 FQAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571

RESULT 8
ABP41875
ID ABP41875 standard; Protein; 408 AA.
XX AC ABP41875;
XX DT 22-AUG-2002 (first entry)
XX DE Human ovarian antigen HISBE12, SEQ ID NO:3007.
XX KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW inflammatory condition; immune disorder; blood disorder;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disorder; urinary system disorder; drug screening;
KW gene therapy; chromosome mapping; forensic analysis;
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW antiinflammatory; gynaecological; reproductive.
XX OS Homo sapiens.
XX PN WO200200677-A1.
XX PD 03-JAN-2002.
XX PF 07-JUN-2001; 2001WO-US18569.
XX PR 07-JUN-2000; 2000US-209467P.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Birse CE, Rosen CA;
XX DR WPI; 2002-147878/19.
XX DR N-PSDB; ABQ54952.
XX PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -
XX PS Claim 11; SEQ ID NO 3007; 2922pp; English.
XX PS
XX CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired

CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders,
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents a human ovarian antigen of the invention.
CC Note: the sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 408 AA;

Query Match 68.9%; Score 2018; DB 23; Length 408;
Best Local Similarity 100.0%; Pred. No. 6.6e-184;
Matches 393; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 RIDQLTNILAPMAVQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAG 238
Db 16 RIDQLTNILAPMAVQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAG 75
QY 239 LKEEETELKQLNLHKDTEPKLEGTHLMGVKDSNIHELEHEQEP TCASQMAEPPRTFRDG 298
Db 76 LKEEETELKQLNLHKDTEPKLEGTHLMGVKDSNIHELEHEQEP TCASQMAEPPRTFRDG 135
QY 299 WVSYYNPVFLAGMGLAPLYMTVLGFCITTYAYTQGLSGSILSLMGASAITGIMGTV 358
Db 136 WVSYYNPVFLAGMGLAPLYMTVLGFCITTYAYTQGLSGSILSLMGASAITGIMGTV 195
QY 359 AFTWLRKCGLVRTLISGLAQLSCLILCVISVMPGSPDLVSPPEDISRFTQGESI 418
Db 196 AFTWLRKCGLVRTLISGLAQLSCLILCVISVMPGSPDLVSPPEDISRFTQGESI 255
QY 419 TPTKIPETITTYMNGNSANIVPETSPEVPIISVLSLLFAGVIAARIGLWSFDLTVTQ 478
Db 256 TPTKIPETITTYMNGNSANIVPETSPEVPIISVLSLLFAGVIAARIGLWSFDLTVTQ 315
QY 479 LLQENVIESERGIINGVQNSMNYLLDLHLHFMVILAPNPEAFGLLVLSVSFVAMGHIMY 538
Db 316 LLQENVIESERGIINGVQNSMNYLLDLHLHFMVILAPNPEAFGLLVLSVSFVAMGHIMY 375
QY 539 FRFAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
Db 376 FRFAQNTLGNKLFACGPDPAKEVRKENQANTSVV 408

RESULT 9

ABB89115
ID ABB89115 standard; Protein; 382 AA.

XX AC ABB89115;

XX DT 24-MAY-2002 (first entry)

XX DE Human polypeptide SEQ ID NO 1491.

XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.

XX OS Homo sapiens.

XX PN WO200190304-A2.

XX PD 29-NOV-2001.

XX PF 18-MAY-2001; 2001WO-US16450.

XX 19-MAY-2000; 2000US-205515P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Birse CE, Rosen CA;
XX WPI; 2002-122018/16.
DR N-PSDB; ABL9524.
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.
XX Claim 11; SEQ ID NO 1491; 2081pp + Sequence Listing; English.
XX The invention relates to novel genes (ABL9449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 382 AA;
Query Match 61.1%; Score 1789; DB 23; Length 382;
Best Local Similarity 92.4%; Pred. No. 4.7e-162;
Matches 353; Conservative 3; Mismatches 26; Indels 0; Gaps 0;
Qy 190 MAVQIMTFGSPVIGCGFISGWNLSVCMVEVLLWKVYQKTPALAVKAGLKEBETELKQL 249
Db 1 MAVQIMTFGSPVIGCGFISGWNLSVCMVEVLLWKVYQKTPALAVKAGLKEBETELKQL 60
Qy 250 NLHKDTEPKPLEGTHLMGVKDSNIHELEHQEPTCASQMAEPFRTRDGMVSYNQPVFL 309
Db 61 NLHKDTEPKPLEGTHLMGVKDSNIHELEHQEPTCASQMAEPFRTRDGMVSYNQPVFL 120
Qy 310 AGMLAFLYMTVLGFDCTTTCYAYTQGLSGSILSILMGASAITGIMGTVAFTWLRKCGL 369
Db 121 AGMLAFLYMTVLGFDCTTTCYAYTQGLSGSILSILMGASAITGIMGTVAFTWLRKCGL 180
Qy 370 VRTGLISGLAQSLCLILCVISVFMPSGPLDLSVSPEDIRSFQGESITPTKIPETITE 429
Db 181 GSAGLISGLAQSLCLILCVISVFMPSGPLDLSVSPEDIRSFQGESITPTKIPETITE 240
Qy 430 IYMSNGSNSANIVPETSPEVPIISVLSLLFAGVIAARIGLWSFDLTVTLQENVIESER 489
Db 241 IYMSNGSNSANIVPETSPEVPIISVLSLLFAGVIAARIGLWSFDLTVTLQENVIESER 300
Qy 490 GIINGVQNSMNYLLDLHLFIMVILAPNPEAFGLLVLSVSVFAMGHIMYFRFAQNTLGNK 549
Db 301 GIINGVQNSMNYLLDLHLFIMVILAPNPEAFGLLVLSVSVFAMGHIMYFRFAQNTLGNK 360
Qy 550 LFACGPDPAKVRKENQANTSV 571
Db 361 LFACGPDPAKVRKENQANTSV 382
RESULT 10
AAW74869

ID XX AAW74869 standard; Protein; 383 AA.
AC XX AAW74869;
XX 25-JAN-1999 (first entry)
DT XX Human secreted protein encoded by gene 141 clone HCMH30.
DE XX
XX Human; secreted protein; testis; tumour; foetal brain tissue;
KW fusion protein; cancer; central nervous system; seizure;
KW diagnosis; neurodegenerative disease.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Misc-difference 383
FT /label= unknown
FT
XX WO9839448-A2.
PN 11-SEP-1998.
PD 06-MAR-1998; 98WO-US04493.
PF 02-OCT-1997; 97US-0061060.
PR 07-MAR-1997; 97US-0038621.
PR 07-MAR-1997; 97US-0040161.
PR 07-MAR-1997; 97US-0040162.
PR 07-MAR-1997; 97US-0040163.
PR 07-MAR-1997; 97US-0040163.
PR 07-MAR-1997; 97US-0040333.
PR 07-MAR-1997; 97US-0040334.
PR 07-MAR-1997; 97US-0040336.
PR 07-MAR-1997; 97US-0040626.
PR 11-APR-1997; 97US-0043311.
PR 11-APR-1997; 97US-0043312.
PR 11-APR-1997; 97US-0043313.
PR 11-APR-1997; 97US-0043314.
PR 11-APR-1997; 97US-0043568.
PR 11-APR-1997; 97US-0043569.
PR 11-APR-1997; 97US-0043576.
PR 11-APR-1997; 97US-0043578.
PR 11-APR-1997; 97US-0043580.
PR 11-APR-1997; 97US-0043669.
PR 11-APR-1997; 97US-0043670.
PR 11-APR-1997; 97US-0043671.
PR 11-APR-1997; 97US-0043672.
PR 11-APR-1997; 97US-0043674.
PR 23-MAY-1997; 97US-0047492.
PR 23-MAY-1997; 97US-0047500.
PR 23-MAY-1997; 97US-0047501.
PR 23-MAY-1997; 97US-0047502.
PR 23-MAY-1997; 97US-0047503.
PR 23-MAY-1997; 97US-0047581.
PR 23-MAY-1997; 97US-0047582.
PR 23-MAY-1997; 97US-0047583.
PR 23-MAY-1997; 97US-0047584.
PR 23-MAY-1997; 97US-0047585.
PR 23-MAY-1997; 97US-0047586.
PR 23-MAY-1997; 97US-0047587.
PR 23-MAY-1997; 97US-0047588.
PR 23-MAY-1997; 97US-0047589.
PR 23-MAY-1997; 97US-0047590.
PR 23-MAY-1997; 97US-0047592.
PR 23-MAY-1997; 97US-0047593.
PR 23-MAY-1997; 97US-0047594.
PR 23-MAY-1997; 97US-0047595.
PR 23-MAY-1997; 97US-0047596.
PR 23-MAY-1997; 97US-0047597.
PR 23-MAY-1997; 97US-0047598.
PR 23-MAY-1997; 97US-0047599.
PR 23-MAY-1997; 97US-0047600.
PR 23-MAY-1997; 97US-0047601.
PR 23-MAY-1997; 97US-0047612.

23-MAY-1997; 97US-0047613.
 23-MAY-1997; 97US-0047614.
 23-MAY-1997; 97US-0047615.
 23-MAY-1997; 97US-0047617.
 23-MAY-1997; 97US-0047618.
 23-MAY-1997; 97US-0047632.
 23-MAY-1997; 97US-0047633.
 06-JUN-1997; 97US-0048964.
 06-JUN-1997; 97US-0048974.
 13-JUN-1997; 97US-0049610.
 08-JUL-1997; 97US-0051926.
 16-JUL-1997; 97US-0052874.
 18-AUG-1997; 97US-0055724.
 22-AUG-1997; 97US-0055630.
 22-AUG-1997; 97US-0056631.
 22-AUG-1997; 97US-0056632.
 22-AUG-1997; 97US-0056636.
 22-AUG-1997; 97US-0056637.
 22-AUG-1997; 97US-0056662.
 22-AUG-1997; 97US-0056664.
 22-AUG-1997; 97US-0056845.
 22-AUG-1997; 97US-0056862.
 22-AUG-1997; 97US-0056864.
 22-AUG-1997; 97US-0056872.
 22-AUG-1997; 97US-0056874.
 22-AUG-1997; 97US-0056875.
 22-AUG-1997; 97US-0056876.
 22-AUG-1997; 97US-0056877.
 22-AUG-1997; 97US-0056878.
 22-AUG-1997; 97US-0056879.
 22-AUG-1997; 97US-0056880.
 22-AUG-1997; 97US-0056881.
 22-AUG-1997; 97US-0056882.
 22-AUG-1997; 97US-0056884.
 22-AUG-1997; 97US-0056886.
 22-AUG-1997; 97US-0056887.
 22-AUG-1997; 97US-0056888.
 22-AUG-1997; 97US-0056889.
 22-AUG-1997; 97US-0056892.
 22-AUG-1997; 97US-0056893.
 22-AUG-1997; 97US-0056894.
 22-AUG-1997; 97US-0056903.
 22-AUG-1997; 97US-0056908.
 22-AUG-1997; 97US-0056909.
 22-AUG-1997; 97US-0056910.
 22-AUG-1997; 97US-0056911.
 05-SEP-1997; 97US-0057650.
 05-SEP-1997; 97US-0057659.
 05-SEP-1997; 97US-0057761.
 12-SEP-1997; 97US-0058785.
 (HUMA-) HUMAN GENOME SCI INC.
 Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
 Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;
 Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
 Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
 WPI; 1998-506364/43.
 N-PSDB; AAV59651.
 New isolated human genes and the secreted polypeptide(s) they encode
 - useful for diagnosis and treatment of e.g. cancers, neurological
 disorders, immune diseases, inflammation or blood disorders
 Claim 1; Page 624-625; 721pp; English.
 This sequence represents a secreted human protein encoded by the nucleic
 acid molecule designated Gene 141 from the human cDNA clone HCM5H30
 (deposited as clone ATCC 97902 and ATCC 209048).
 The gene can be used to generate fusion proteins by linking to the gene
 to a human immunoglobulin Fc portion (e.g. AAV59502) for increasing the
 stability of the fused protein as compared to the human protein only.

CC The invention relates to 186 novel genes and their fragments (nucleic
 CC acid sequences: AAV59511-V59812; amino acid sequences AA74731-W75026)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 186
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAV59511 for described uses).
 XX

SQ Sequence 383 AA;

Query Match 61.1%; Score 1789; DB 19; Length 383;

Best Local Similarity 92.4%; Pred. No. 4.7e-162;

Matches 353; Conservative 3; Mismatches 26; Indels 0; Gaps 0;

QY 190 MAVQIMTFGSPVIGCGFISGNILVSMCEYVLLWKVYQKTPALAVKAGLKEEETELKQL 249

DB 1 MAVQIMTFGSPVIGCGFISGNILVSMCEYVLLWKVYQKTPALAVKAGLKEEETELKQL 60

QY 250 NLHKDTEPKLEGTHLMGVKDSNIHELEHEOEPTCASQMAEPFRFRDGVSYYNQPVFL 309

DB 61 NLHKDTEPKLEGTHLMGVKDSNIHELEHEOEPTCASQMAEPFRFRDGVSYYNQPVFL 120

QY 310 AGMGLAFLYMTVLGFDCTTTGYAYTQGLSGSILSGASAITGIMGTVAFTLRRKCGL 369

DB 121 AGMGLAFLYMTVLGFDCTTTGYAYTQGLSGFHPQYFDGSIYNNWNGCSFYLATSKWVF 180

QY 370 VRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPFEDIRSRFIQGESITPKIPEITTE 429

DB 181 GSAGLISGLAQLSCLILCVISVFMPSGLDLSVSPFEDIRSRFIQGESITPKIPEITTE 240

QY 430 IYMSNGSNSANIVPETSPEVPIISVLLFAGVIAARIGLWSFDLTVTQLLOENVIESER 489

DB 241 IYMSNGSNSANIVPETSPEVPIISVLLFAGVIAARIGLWSFDLTVTQLLOENVIESER 300

QY 490 GIINGVQNSMNYLLDLHFIWVILAPNPEAFGLLLVLSVSVFVAMGHIMYFRFAQNTLGNK 549

DB 301 GIINGVQNSMNYLLDLHFIWVILAPNPEAFGLLLVLSVSVFVAMGHIMYFRFAQNTLGNK 360

QY 550 LFACGPDPAKEVRKENQANTSVV 571

DB 361 LFACGPDPAKEVRKENQANTSVV 382

RESULT 11

ABG95320

ID ABG95320 standard; Protein; 383 AA.

AC ABG95320;

DT 15-JAN-2003 (first entry)

DE Human novel secreted protein #141.

DE Human; secreted protein; autoimmune disease; chemotaxis;

KW rheumatoid arthritis; hyperproliferative disorder; breast neoplasm;

KW liver neoplasm cardiovascular disorder; cardiac arrest; skin aging;

KW cerebrovascular disorder; cerebral ischaemia; angiogenesis; sunburn;

KW nervous system disorders; Alzheimer's disease; infection;

KW ocular disorder; corneal infection; wound healing; tissue regeneration;

KW epithelial cell proliferation; organ transplantation; food additive;

KW preservative; nutritional.

OS Homo sapiens.

XX US6420526-B1.

PN 16-JUL-2002.

PD 08-SEP-1998; 98US-0149476.

XX 07-MAR-1997; 97US-038621P.

QY 190 MAVGQIMTFGSPVIGCGFISGNLVMCMVEYLLMKVYQKTPALAVKAGLKEEETELKQL 249
DB 1 MAVGQIMTFGSPVIGCGFISGNLVMCMVEYLLMKVYQKTPALAVKAGLKEEETELKQL 60
QY 250 NLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPFRFRDQWVSYNQPVFL 309
DB 61 NLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPFRFRDQWVSYNQPVFL 120
QY 310 AGMGLAFYMTVLGFDCTTGVAYTQGLSGSILSLMGASAITGIMGTVAFTWLRKGL 369
DB 121 AGMGLAFYMTVLGFDCTTGVAYTQGLSGSILSLMGASAITGIMGTVAFTWLRKGL 180
QY 370 VRTGLISGLAQSLCILCVISVFMFGSPDLDSVPFEDIRSRFQGESITPKIPEITTE 429
DB 181 GSAGLISGLAQSLCILCVISVFMFGSPDLDSVPFEDIRSRFQGESITPKIPEITTE 240
QY 430 IYMSGNSANIVPETSPEVPIISVLLFAGVIAARIGLSFDTVTQLQENVIBSER 489
DB 241 IYMSGNSANIVPETSPEVPIISVLLFAGVIAARIGLSFDTVTQLQENVIBSER 300
QY 490 GIINGVQNSMNYLLDLHFIMVILAPNEAFGLLVLSVFMVAMGHIMYFPAQNTLGNK 549
DB 301 GIINGVQNSMNYLLDLHFIMVILAPNEAFGLLVLSVFMVAMGHIMYFPAQNTLGNK 360
QY 550 LFACGPDAAKEVRKENQANTSVV 571
DB 361 LFACGPDAAKEVRKENQANTSVV 382

RESULT 12
AAM40920

ID AAM40920 standard; Protein; 439 AA.

AC AAM40920;

XX 22-OCT-2001 (first entry)

DT Human polypeptide SEQ ID NO 5851.

DE Human; neutropic; immunosuppressant; cytostatic; gene therapy; cancer;

XX peripheral nervous system; neuropathy; central nervous system; CNS;

KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;

KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;

KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;

KW leukaemia.

XX Homo sapiens.

OS WO200153312-A1.

XX 26-JUL-2001.

PD 26-DEC-2000; 2000WO-US34263.

PF 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-0552317.

PR 09-JUL-2000; 2000US-0598042.

PR 19-JUL-2000; 2000US-0620312.

PR 03-AUG-2000; 2000US-0653450.

PR 14-SEP-2000; 2000US-0662191.

PR 19-OCT-2000; 2000US-0693036.

PR 29-NOV-2000; 2000US-0727344.

XX (HYSEQ-) HYSEQ INC.

DA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

PI Zhao QA, Zhou P, Goodrich R, Drmanac RT,

XX WPI; 2001-442253/47.

DR N-PSDB; AAI60076.

XX

PT Novel nucleic acids and polypeptides, useful for treating disorders
XX such as central nervous system injuries -
PS Example 2; SEQ ID NO 5851; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC the encoded polypeptides (AAM38642-AAM42213) with neutropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.

XX Sequence 439 AA;

Query Match 55.7%; Score 1630; DB 22; Length 439;

Best Local Similarity 78.9%; Pred. No. 9.3e-147;

Matches 333; Conservative 8; Mismatches 31; Indels 50; Gaps 6;

QY 1 MTRAGDHNRRGCGGSLADYLTSAKFLLYLGHSLSTWGRMWHFAVSVFLVELYGNLSLL 60

DB 35 MTRAGDHNRRGCGGSLADYLTSAKFLLYLGHSLSTWGRMWHFAVSVFLVELYGNLSLL 94

QY 61 TAVYGLVWAGSVLVGLGAIIGDWVDKARLKVACTSLVQNVSVLTCGIIIMVFLKHHEL 120

DB 95 TAVYGLVWAGSVLVGLGAIIGDWVDKARLKVACTSLVQNVSVLTCGIIIMVFLKHHEL 154

QY 121 LTMVHGWLTVSCYILITITIANIANLASTATAITQRDWIVVWAGEDRSKLANNNATIRRI 180

DB 155 LTMVHGWLTVSCYILITITIANIANLASTATAITQRDWIVVWAGEDRSKLANNNATIRRI 214

QY 181 DQLTNLAPMAVGQIMTFGSPVIGCGFISGNLVMCMVEYLLMKVYQKTPALAVKAGLK 240

DB 215 DQLTNLAPMAVGQIMTFGSPVIGCGFISGNLVMCMVEYLLMKVYQKTPALAVKAGLK 274

QY 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPFRFRDQW 300

DB 275 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPFRFRDQW 334

QY 301 SYYNQPVFLAGMGLAF-LYMTVLGFDCTTGVAYTQGLSGSILSLM 346

DB 335 SYYNQPVFLAGMGLAF-LYMTVLGFDCTTGVAYTQGLSGSILSLM 383

QY 347 GASAITGIMGTVAFTWLRKRC-----GLVRTGLISGLAQSLCLII-LCVISVFM 393

DB 384 -----SYNNWNGNSPYLATSQWFGSDRLRIGTAFGLFDVLCDCIHAWK 431

QY 394 PG 395

DB 432 PG 433

RESULT 13

ABG20262

ID ABG20262 standard; Protein; 439 AA.

XX AC ABG20262;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #20253.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

KW

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 18:30:57 ; Search time 23 Seconds
(without alignments)
1050.413 Million cell updates/sec

Title: US-09-715-927-6

Perfect score: 2929

Sequence: 1 MTRAGDHNRQCCGSLADY.....ACGPDAKEVRKENQANTSVW 571

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA*

1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/ECTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1789	61.1	383	4	US-09-149-476-460
2	123.5	4.2	524	2	US-08-928-632-12
3	123.5	4.2	524	4	US-09-339-972-12
4	112	3.8	305	4	US-09-252-991A-33032
5	110	3.8	429	4	US-09-328-352-4392
6	109.5	3.7	470	4	US-09-134-001C-4610
7	108	3.7	330	1	US-08-118-270-19
8	108	3.7	330	5	PCT-US93-08528-19
9	108	3.7	618	4	US-08-595-553A-2
10	108	3.7	618	4	US-09-640-198D-4
11	107	3.7	411	4	US-09-134-001C-3299
12	107	3.7	465	4	US-09-328-352-5222
13	104	3.6	534	2	US-09-031-332-4
14	104	3.6	534	3	US-09-299-549-4
15	104	3.6	534	4	US-09-610-417-4
16	103.5	3.5	471	4	US-09-328-314-17
17	103	3.5	358	2	US-08-748-485-6
18	102.5	3.5	529	4	US-09-291-922-28
19	102	3.5	553	4	US-09-134-001C-2974
20	101.5	3.5	996	4	US-09-252-991A-27018
21	100.5	3.4	411	3	US-09-236-080-6
22	100.5	3.4	760	4	US-09-252-991A-27790
23	100	3.4	499	4	US-09-134-001C-5370
24	99.5	3.4	461	1	US-08-194-338-4
25	98.5	3.4	411	3	US-09-236-080-2
26	98.5	3.4	411	4	US-09-336-643A-83
27	98	3.3	366	4	US-09-134-001C-5502

28	98	3.3	443	4	US-09-134-001C-3227	Sequence 3227, Ap
29	98	3.3	808	4	US-09-134-001C-3105	Sequence 3105, Ap
30	97.5	3.3	596	4	US-09-252-991A-26031	Sequence 26031, A
31	97.5	3.3	808	2	US-08-629-231A-33	Sequence 33, Appl
32	97.5	3.3	808	2	US-08-658-335B-33	Sequence 33, Appl
33	97.5	3.3	808	4	US-09-406-640-33	Sequence 20611, A
34	97.5	3.3	1033	4	US-09-252-991A-20611	Sequence 4, Appli
35	97	3.3	359	2	US-08-103-170-4	Sequence 6, Appli
36	97	3.3	359	2	US-08-103-170-6	Sequence 6, Appli
37	97	3.3	457	2	US-08-882-704A-6	Sequence 6, Appli
38	97	3.3	457	4	US-09-151-957-6	Sequence 17, Appl
39	97	3.3	484	4	US-09-266-965-17	Sequence 5530, Ap
40	96	3.3	427	4	US-09-107-532A-5530	Sequence 2, Appli
41	96	3.3	1527	4	US-09-376-330-2	Sequence 8, Appli
42	95.5	3.3	370	4	US-09-144-914-8	Sequence 6160, Ap
43	95.5	3.3	480	4	US-09-107-532A-6160	Sequence 17805, A
44	95	3.2	1421	4	US-09-252-991A-17805	Sequence 4353, Ap
45	94.5	3.2	398	4	US-09-134-001C-4353	

ALIGNMENTS

RESULT 1

US-09-149-476-460
; Sequence 460, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23


```
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

Query Match      61.1%; Score 1789; DB 4; Length 383;
Best Local Similarity 92.4%; Pred. No. 8.8e-175;
Matches 353; Conservative 3; Mismatches 26; Indels 0; Gaps 0;

QY 190 MAVGQIMTFGSPVTCGCGFISGNLVSMCVYVLLWKVYQKTPALAVKAGLKEEETELKQL 249
Db 1 MAVGQIMTFGSPVTCGCGFISGNLVSMCVYVLLWKVYQKTPALAVKAGLKEEETELKQL 60

QY 250 NLHKDTEPKLEGTHLMGVKDSNIHEHEOEPTCASQMAEPFRFRDGMVSYYNQPVFL 309
Db 61 NLHKDTEPKLEGTHLMGVKDSNIHEHEOEPTCASQMAEPFRFRDGMVSYYNQPVFL 120

QY 310 AGMGLAFYIMVTLGPDCTITGAYTQGLSGSILSLMGASAITGIMGTVAFTWLRKRCGL 369
Db 121 AGMGLAFYIMVTLGPDCTITGAYTQGLSGFHPQVFDGSIYNNWNGCSFVLATSKMWF 180

QY 370 VRTGLISGLAQLSCLILCVISVFMFGSPDLDSVSPFEDIRSRFIOGESITTKIPEITTE 429
Db 181 GSAGLISGLAQLSCLILCVISVFMFGSPDLDSVSPFEDIRSRFIOGESITTKIPEITTE 240

QY 430 IYMSGNSANSANIVPTSPESPIISVLLFAGVIAARIGLWSFDLTVTQLLOENVIESER 489
Db 241 IYMSGNSANSANIVPTSPESPIISVLLFAGVIAARIGLWSFDLTVTQLLOENVIESER 300

QY 490 GIINGVQNSMYLLDLHLFIMVILAPNPEAFGLLVLSVSVFAMGHIMYFRFAQNTLGNK 549
Db 301 GIINGVQNSMYLLDLHLFIMVILAPNPEAFGLLVLSVSVFAMGHIMYFRFAQNTLGNK 360

QY 550 LFACGPDKEVRKENQANTSVV 571
Db 361 LFACGPDKEVRKENQANTSVV 382

RESULT 2
US-08-928-692-12
; Sequence 12, Application US/08928692
; Patent No. 5958727
; GENERAL INFORMATION:
; APPLICANT: Brody, Howard
; APPLICANT: Yaver, Deborah S.
; APPLICANT: Lamsa, Michael
; APPLICANT: Hansen, Kim
; TITLE OF INVENTION: Methods for Modifying the Production of
; TITLE OF INVENTION: a Polypeptide
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5958727o No. 5958727disk of No. 5958727th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,692
; FILING DATE: 12-SEPT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4944,200-US
; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 524 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5958727e
US-08-928-692-12

Query Match      4.2%; Score 123.5; DB 2; Length 524;
Best Local Similarity 20.3%; Pred. No. 0.00084;
Matches 99; Conservative 69; Mismatches 150; Indels 169; Gaps 23;

QY 41 MWFHFAVSVELV-----ELYGNSL--LITAVYGLVVGAGVLVGLGAIIGDWDKRNARKLVAQ 93
Db 96 LMSUSVSFAVGVGMTASFFGWLJGDTLGRKAMLVANILSLVGALLMGF-----SKLGP 149

QY 94 TSLVY--QNVSVILCGIILMMVFLHKLHLLTMVHGWLVTSCYILIIITIANIANLASTAT 150
Db 150 SHILIIAGRSISGLYCGLI-----SGLVPMY-----IGEIAPTA- 183

QY 151 AITIQDWMVVVAGEDRSKLANMNATIRIDQLTNILAPMAVGQIMTFGSPVCGFTSG 210
Db 184 -----WNL-----VSMCVYVLLWKVYQKTPALAVKAGLKEEETELKQLNLHKDTEPKPL 260

QY 214 NYDLWHILLGLSGVRAIRLQSLLLFFCPSPRYLYIKL---DEEVKAKQ-SLKR----- 262
Db 261 EGTLMGVKD--SNIHEHEOEPTCASQMAEPFRFRDGMVSYYNQPVFLAGM-GLAFL 317

QY 263 ----LRGYDDVTKDINEWRKEREASSEQKVIITLFTN---SSYRQDILVALMLHVAQQ 315
Db 318 YMTVLGFDCTITGAYTQGLSGSILSLMGASAITGIMGTVAFTWLRKRCGLVRTGLISG 377

QY 316 FSGINGIFYYSTISFTAGISKPVYATI-GVGAVNMVF-TAVSVFLVEKAGRRSLFLIGM 373
Db 378 LAQLSCLI-----LCVISVFMFGSPDLDSVSPFEDIRSRFIOGESITP 420

QY 374 SGMFVCAIFMSVGLVLLNKFMSMSYVSMIAIF-----LFVSFFE-----IGP 415
Db 421 TKIPEITTEIYMSGNSANSANIVPTSPESPIIS-----PRPALATAAFSNWTCNFCVALCFQYIADFCGPYVFFL 468

QY 459 FAGVIAA 465
Db 469 FAGVLLA 475

RESULT 3
US-09-339-372-12
; Sequence 12, Application US/09339972
; Patent No. 6323002
; GENERAL INFORMATION:
; APPLICANT: Brody, Howard
; APPLICANT: Yaver, Deborah S.
; APPLICANT: Lamsa, Michael
; APPLICANT: Hansen, Kim
; TITLE OF INVENTION: Methods for Modifying the Production of
; TITLE OF INVENTION: a Polypeptide
; NUMBER OF SEQUENCES: 80
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6323002o No. 6323002disk of No. 6323002th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```


Db 4 YIDIIRYFYNDLSHSYNFLLTMTYKTSQSFOSLIYM-LIGSAIILALSIGVRHGFGLYL 62
QY 63 -----VGLVAGSVLVLCAL---IGDWVDKNARKVAQTSLVQVNSVILCG 107
Db 63 VPMSEFGWGHVFSLAIAQNLWGAQOPFGAIADKYG-----SKVVTVVGGLLYT 115
QY 108 I-ILMMVFLKHLLTMTYHGWVLTSCYIIITITIANIANLASTATAITORDWVWVAGED 166
Db 116 LGLLLMAF-----SSSVLILNLSGLIIGALSATSFTVLLSAVGRAAPPE 161
QY 167 RSKLANMATIRRIDQLNILAPMAVGQIMFSPVIGCGFISGNW-----LVSMCV 218
Db 162 KRSMAWGIAS-----AAGSGQFIMLPSTLLLLKTV-GWSSALMVSALLIALII 209
QY 219 EYVLLKVVYQKTPALAVKAGLKEETELKQNLNKHDTPEKP-----LEG 262
Db 210 P--LAWML--KGPSNQTPKAIAQPOLTPKQV-LHARKHKEPFWLALGLVCGQVFLG 264
QY 263 THLMGVKDSNIHELEHEOEPCTAS-----QMAEPFRTRDGVW--SYYNQPVFLAGM--- 312
Db 265 VHLPG-----YLIDHGFDAITGTVFLALVGLFNIVGTYGAGWLGDRFSKPKLLMALYGS 318
QY 313 -GLAFLYMTVLGFDCTITGYAY-----TQGLSGSILSILMGASAITGIMGTVAFT 361
Db 319 RGIAIAFLPLSTYTV-YAFGIIMGLLWSTVPLTNGIVANMGVVKYLSMISGIVEFT 377
QY 362 WLRKCGLVRTGLISGL-----AQLSCLILCVISVFMFG 395
Db 378 ---HQVGSFPGWGLGVNHDLTGNNYNAIWLCSIALSIIG 413

RESULT 6

US-09-134-001C-4610
; Sequence 4610, Application US/09134001C
; Patent No. 6380370

; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 4610

; LENGTH: 470

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis

US-09-134-001C-4610

Query Match 3.7%; Score 109.5; DB 4; Length 470;
Best Local Similarity 19.5%; Pred. No. 0.019;

Matches 100; Conservative 85; Mismatches 190; Indels 137; Gaps 21;

QY 73 LVLAGAIGDWVD-----KNARKVAQTSLVQVNSVILCGIILMMVFL 115
Db 45 LILGTIFLDAYDITLGMTDOLTOQPHLSFATISVNTSLPIGALFALLGGTLYOFG 104
QY 116 HKHELLTYHGWVLTSCYIIITIANI-ANLASTATAITIORDWIIVVAGEDR----- 167
Db 105 RKR-----ILSIALLTITVSLGAALAPNVIIICRCIMGFAGNDSVPAFTFI 154
QY 168 SKLANMATIRRID--QNTNLAIPMAVGQIMFSPVIGCG-----FISGNLVSMCVVEY 220
Db 155 AEISNLKHGRNVNVQVWVVAIVTSALVVT-AFFMLGAGHLWRVAVGFGALIAFVLY 213
QY 221 VLLMKVYQKTPALAVKAGLKEEETELKQNLNKHDTPEKPLEGTHLMGVKDSNIHELEHQ 280
Db 214 ILRIKYLHESPTWINHYLSLEKATEFIRKYYHKD-----IHLEGTLEDLNL----- 259

QY 281 EPTCASQMAEPFRTRDGVWVSYYNQPVFLA-----GMGLAFLYMTVLGFDCTIT 328
Db 260 -----SDVTSPHNSWTDLFPKRYIKRIILATAISTALQGMQYVGVG---LYIPIIA----- 306
QY 329 TGYAYTOGLSSILSILMGASA--ITGIMGTVAFTWLRKCGLVRTGLISGLAQLSCLIL 386
Db 307 -----TVLISKDKTGILLGTAVNIAIGTLGAYLAQLTLYKLGTRKLTMI-GFTLVLLSMV 360
QY 387 CVISVFMFGSPDLDSVPFFEDIRSRFTQGESITP-TKIPEITTEIYMSNGSNSANIVPET 445
Db 361 CV-GLFVHHLPMNLNTP-----LIGLFLHSGSGGTQGTCTICALSFPHLSQATGFVES 415
QY 446 SPESVPIISVLLFAGVIAARIIGLWSFDLTVTQLLENVIESBERGIINGVQNSMNYLLDL 505
Db 416 VSRITSGIIGTFVF--PIILAAVGLTN----- 439
QY 506 LHFTMVLAPNPEAFGLLVLSVSFVAMG-HI 536
Db 440 ---TMLILSIVP-LLGIIITVSIKWEAVGRHI 467

RESULT 7

US-08-118-270-19

; Sequence 19, Application US/08118270

; Patent No. 5508384

; GENERAL INFORMATION:

; APPLICANT: Murphy, Randall B.

; APPLICANT: Schuster, David I.

; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN

; NUMBER OF SEQUENCES: 348

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BROWDY AND NEIMARK

; STREET: 419 Seventh Street, N.W., Suite 300

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20004

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/118,270

; FILING DATE: 09-SEP-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/943,236

; FILING DATE: 10-SEP-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Townsend, Kevin G.

; REGISTRATION NUMBER: 34,033

; REFERENCE/DOCKET NUMBER: MURPHY=2A

; TELEPHONE: 202-628-5197

; TELEFAX: 202-737-3528

; TELEX: 248633

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 330 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-08-118-270-19

Query Match 3.7%; Score 108; DB 1; Length 330;
Best Local Similarity 24.4%; Pred. No. 0.016;

Matches 51; Conservative 39; Mismatches 71; Indels 48; Gaps 11;

QY 60 LTAIVGYLV-----VAGSVLVLAGIIGDWVDKNARKVAQTSLVQVNSVILCGIILMMVFL 115
Db 4 LAAVVGFLIVTVVGVNLVVIIVL-----TSRALRAPONLFLVSIASADILVATLVMPFS 58

Db 301 --INVVYKCDPILLTGRISAPDQYMPLLVLDIFEDLPGVGLFLACAYSGTILSTASTS 358
Qy 265 LMVGKDSNIHEHEQEPTCASQMAEPFRTERFDGWVSYNQPVFLAGMGLAFLYMTVLGF 324
Db 359 INAMAATVEDLIKPRMPGLAPR-----KLVFIS-KGLSFIY---GS 396
Qy 325 DCITTYAYTOGLSGSILMGLASAITG---IMGTV-----AFT--WLRRKQGL--VR 371
Db 397 ACLTV-----AALSSLLGGVQLQGSFTVMGVISGPLIGAFITLGMLLPACNTPGVL 446
Qy 372 TGLISGLAQLSCLILCVISVMPGSPDLDSVPFDIRSRFIQGESITPTKIPBITTEIY 431
Db 447 SGLAAGLA-VSLWAVGATILPPGE-----QTMGVLP TSAAGCTNDSV 488
Qy 432 M---SNGSNANIVPETSPE-----SVPIISVLLFAGVIAARIGLWSPDLTWTQLL 480
Db 489 LLGPPGATNASGIPSSGMDTGRPALADTFVAYSILYVYGALG-----TLTMTL 536
RESULT 10
US-09-640-198D-4
; Sequence 4, Application US/09640198D
; Patent No. 6586411
; GENERAL INFORMATION:
; APPLICANT: Russell, Stephen
; APPLICANT: Kay Whye, Peng
; TITLE OF INVENTION: System for Monitoring the Location of
; FILE REFERENCE: 07039-295001
; CURRENT APPLICATION NUMBER: US/09/640,198D
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 60/149,168
; PRIOR FILING DATE: 1999-08-17
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-640-198D-4

Query Match 3.7%; Score 108; DB 4; Length 618;
Best Local Similarity 19.1%; Pred. No. 0.042;
Matches 103; Conservative 71; Mismatches 160; Indels 204; Gaps 27;
Qy 14 CGSLADYLSAKFLYLHLSLSTWGDMMHFAVSFLVELYNSILLTAVYGLVAVGSVL 73
Db 132 CGTL-QYLVAT--MLYTG-----IVIYAPALILNQVTGLDIWASLL 169
Qy 74 VLGAIGDWDKNAKLKVAQTSVLVQNVSVILCGIILMMVFLKHHELLTMYHGWV-LTSC 132
Db 170 STGLICTLYTVGKKAVVWTD--VFQVVMVVG-----WVILARG 209
Qy 133 YLIIITIANIANLASTATAITQ-----RDWIVVVG-----164
Db 210 VILGGPRNVLSLAQNSHRSINLMDFPDPRSRYTFWTFIVGTLVLSWYGVNQVQRY 269
Qy 165 -----EDRSKLANMNATRRIDQTLNLA PMAVGQIMTFGSPVGGCFISGNLVSMCVE 219
Db 270 VACHTEGKAKLALL-----VNQL-----GLFLIVASAAC-CG-----300
Qy 220 YVLLMKVYOK--TPALAVKAGLKEEETELKOLNHLKDTPEKP-----LEGTH 264
Db 301 --IVMFYKCDPILLTGRISAPQYMPLLVLDIFEDLPGVGLFLACAYSGTILSTASTS 358
Qy 265 LMVGKDSNIHEHEQEPTCASQMAEPFRTERFDGWVSYNQPVFLAGMGLAFLYMTVLGF 324
Db 359 INAMAATVEDLIKPRMPGLAPR-----KLVFIS-KGLSFIY---GS 396
Qy 325 DCITTYAYTOGLSGSILMGLASAITG---IMGTV-----AFT--WLRRKQGL--VR 371
Db 397 ACLTV-----AALSSLLGGVQLQGSFTVMGVISGPLIGAFITLGMLLPACNTPGVL 446

RESULT 12

Qy 372 TGLISGLAQLSCLILCVISVMPGSPDLDSVPFDIRSRFIQGESITPTKIPBITTEIY 431
Db 447 SGLAAGLA-VSLWAVGATILPPGE-----QTMGVLP TSAAGCTNDSV 488
Qy 432 M---SNGSNANIVPETSPE-----SVPIISVLLFAGVIAARIGLWSPDLTWTQLL 480
Db 489 LLGPPGATNASGIPSSGMDTGRPALADTFVAYSILYVYGALG-----TLTMTL 536
RESULT 11
US-09-134-001C-3299
; Sequence 3299, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3299
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3299

Query Match 3.7%; Score 107; DB 4; Length 411;
Best Local Similarity 19.7%; Pred. No. 0.028;
Matches 96; Conservative 70; Mismatches 167; Indels 154; Gaps 20;
Qy 17 LADYLTSAKFLYLHLSLSTWGDMMHFAVSFLVELYNS-----LLLTAV 63
Db 19 MAKYFPSSFLFLGNWIGVG-----LWVFLTYHNNAVYGLVNFCLIFILLLSV 71
Qy 64 YGLVAGSVLVGALIGDWDKNAKLKVAQTSVLVQNVSVILCGIILMMVFLKHHELLTM 123
Db 72 W-----AGS-----IADKYDKGLLRITITSSFI--ITALLC-----LLT- 104
Qy 124 YHGWLTSCYILII---TIANIANLASTATAITQIRDWIVVWVAGEDRSKLANMNATIRR 179
Db 105 ---YLSNIIIFILLYATFRGILSAVETPVRAVLPDL-----SSKISTQAVSFH 153
Qy 180 ---IDQLTNILAPMAVGQIMTFGSPV-----ICGCFISGNLVSMCVYVLLMKVYQKTP 231
Db 154 SFIINICRSIGPALAGGLIAVYHTPTTFLAQAVCYFIAVLCIPHFVILSKEGKALP 213
Qy 232 ALAVKAGLKEEETELKOLNHLKDTPEKPLEGTHLMGVKOSNIHELEHEQBP TCASQMAEP 291
Db 214 LKV-----LNYFKSN-----LEGSQ-----229
Qy 292 FRFRDGVSYNQPVFGLAG---MGLAFLYMTVLGPDCTITTYAYTOGLSGSILMGLA 348
Db 230 -----IFITSIIIMATGFSYTVLP---VLTNHFPG--QSQVFGIAMTF 269
Qy 349 SAITGIMGTVAFTWLRRKQGLVR-----TGLISGLAQLSCLILCVISVMPGSPDLDSVSP 404
Db 270 CAIGGIVATIVLPSILKHLSTVKMYLSSILFGIALLGIIHHLVVMFCITLIGL-FSQ 328
Qy 405 FEDIRSRFIQGESITPTKIPBITTEIYMSNG-----SNSANIVPETSPEVPI 452
Db 329 WARTNRVYFQHSVKDCRGKVLIIIMDRGMIPGLSLIMSFFADMEGILTTFTIMGIST 388
Qy 453 ISVSLLF 459
Db 389 ISISLIF 395

STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/299,549
FILING DATE: 26-APR-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/031,392
FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Meiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 534 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-299-549-4

Query Match 3.6%; Score 104; DB 3; Length 534;
Best Local Similarity 20.1%; Pred. No. 0.086;
Matches 89; Conservative 73; Mismatches 126; Indels 154; Gaps 24;
Qy 42 WHPAVSVFLVLYGNSLLLTAVYGLVAGSVLVGAIIGDWV-DKNARKLVAQTSLLVQV 100
Db 109 WLSVSMF-----AVGGMVSSFTV-----GWIGDLGRVK-----AMLVN 144
Qy 101 VSVILCGIILMMVFLHKKHELLTMYHGVLTSCYILITIANIANLASTATAITQDVI 160
Db 145 VLSIAGNLLMGLAKMGPSHIL-IIAGRAITGLY-----CGLSSGLVP-----M 186
Qy 161 VVAGEDSKLANMNAIRRIDQLTNLAIPMAVQIMTFGSPVIGCGFISG-----WNL--- 213
Db 187 YVSEVSPALRGALGTLHQAIVTGILI-----SQVLGLDFLLGNDELWPLLLG 235
Qy 214 ---VSMCEVYVLWKVYQKTPALAVKAG-LKEBETELKQLNLKHDTEPKPLEGTHLMGVK 269
Db 236 LSGVAALLQFLLLLCPESPRYLYIKLGKVEAKSKLRGNC-----PMK----- 283
Qy 270 DSNIEHEHEQEPCTAS-----QMAEPRTFRDGMVSYNQ 305
Db 284 --EIAEMEKEQEAASEKRVISIGLFSKKYRQAVIVALMVQISQQFSGINA--IFYST 339
Qy 306 PVF-LAGMGLAFLYMTVLGPDCTTGYAYTQGLSGSILSLMGASA-----ITGIMG-- 356
Db 340 NIFQAGVGQPVYIYATI-GVGVNVTVF-----TVISVFLVEKAGRRSLFLAGLMGL 390
Qy 357 --TVAFWLRRKCGLVRTGLISGLAQSLCILCVISVFM-----PG----- 395
Db 391 ISAVAMT-----VGLV-----LLSQFAMSVYVSMVAIFLVIFFEVFGPGPIPWFI 442
Qy 396 --SPLDLSVSPFDIRSRFIQ 415
Db 443 GPRPAATAVAGFCNACNFI 464

RESULT 15

US-09-610-417-4
Sequence 4, Application US/09610417
Patent No. 6346374
GENERAL INFORMATION:
APPLICANT: Tartaglia, Louis A.

Weng, Xun
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
GLUTEX AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/610,417
FILING DATE: 05-Jul-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/299,549
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Meiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 534 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-610-417-4

Query Match 3.6%; Score 104; DB 4; Length 534;
Best Local Similarity 20.1%; Pred. No. 0.086;
Matches 89; Conservative 73; Mismatches 126; Indels 154; Gaps 24;
Qy 42 WHPAVSVFLVLYGNSLLLTAVYGLVAGSVLVGAIIGDWV-DKNARKLVAQTSLLVQV 100
Db 109 WLSVSMF-----AVGGMVSSFTV-----GWIGDLGRVK-----AMLVN 144
Qy 101 VSVILCGIILMMVFLHKKHELLTMYHGVLTSCYILITIANIANLASTATAITQDVI 160
Db 145 VLSIAGNLLMGLAKMGPSHIL-IIAGRAITGLY-----CGLSSGLVP-----M 186
Qy 161 VVAGEDSKLANMNAIRRIDQLTNLAIPMAVQIMTFGSPVIGCGFISG-----WNL--- 213
Db 187 YVSEVSPALRGALGTLHQAIVTGILI-----SQVLGLDFLLGNDELWPLLLG 235
Qy 214 ---VSMCEVYVLWKVYQKTPALAVKAG-LKEBETELKQLNLKHDTEPKPLEGTHLMGVK 269
Db 236 LSGVAALLQFLLLLCPESPRYLYIKLGKVEAKSKLRGNC-----PMK----- 283
Qy 270 DSNIEHEHEQEPCTAS-----QMAEPRTFRDGMVSYNQ 305
Db 284 --EIAEMEKEQEAASEKRVISIGLFSKKYRQAVIVALMVQISQQFSGINA--IFYST 339
Qy 306 PVF-LAGMGLAFLYMTVLGPDCTTGYAYTQGLSGSILSLMGASA-----ITGIMG-- 356
Db 340 NIFQAGVGQPVYIYATI-GVGVNVTVF-----TVISVFLVEKAGRRSLFLAGLMGL 390
Qy 357 --TVAFWLRRKCGLVRTGLISGLAQSLCILCVISVFM-----PG----- 395
Db 391 ISAVAMT-----VGLV-----LLSQFAMSVYVSMVAIFLVIFFEVFGPGPIPWFI 442
Qy 396 --SPLDLSVSPFDIRSRFIQ 415

Db 443 GPRPAIAVAGFCNWACNFIVG 464

Search completed: November 5, 2003, 18:37:14
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 18:32:22 ; Search time 45 Seconds
(without alignments)
2179.308 Million cell updates/sec

Title: US-09-715-927-6
Perfect score: 2929
Sequence: 1 MTRAGDHNRRGCCGSLADY.....ACGPDAKEVRKENQANTSW 571

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2924	99.8	571	11	US-09-973-180-2
2	2924	99.8	571	11	US-09-973-180-3
3	2922	99.8	571	11	US-09-973-180-4
4	2917	99.6	571	9	US-09-729-674-130
5	1789	61.1	383	11	US-09-809-391-460
6	1789	61.1	383	12	US-09-882-171-460
7	125.5	4.3	575	15	US-10-156-761-10350
8	120	4.1	458	12	US-10-283-423-6
9	120	4.1	458	12	US-10-213-821-6
10	119	4.1	537	10	US-09-738-626-6730
11	114.5	3.9	424	10	US-09-738-626-3641
12	114	3.9	392	12	US-09-769-744A-120
13	113	3.9	522	9	US-09-804-551B-38
14	108.5	3.7	486	12	US-10-419-190-2
15	108	3.7	618	10	US-09-995-007-2

16	107.5	3.7	488	9	US-09-815-242-10767	Sequence 10767, A
17	104	3.6	534	10	US-09-981-947A-4	Sequence 4, Appli
18	104	3.6	608	15	US-10-156-761-14861	Sequence 14861, A
19	103.5	3.5	471	9	US-09-989-861-17	Sequence 17, Appl
20	102.5	3.5	473	10	US-09-746-491-72	Sequence 72, Appl
21	102.5	3.5	529	14	US-10-051-902-28	Sequence 28, Appl
22	102.5	3.5	529	14	US-10-051-909-28	Sequence 28, Appl
23	101.5	3.5	426	15	US-10-187-267A-43	Sequence 43, Appl
24	101.5	3.5	435	15	US-10-156-761-11512	Sequence 11512, A
25	101.5	3.5	739	12	US-10-328-194A-3	Sequence 3, Appli
26	101.5	3.5	739	15	US-10-305-823-369	Sequence 369, App
27	101	3.4	455	9	US-09-815-242-13794	Sequence 13794, A
28	101	3.4	551	14	US-10-216-355-4	Sequence 4, Appli
29	100.5	3.4	411	9	US-09-828-746-6	Sequence 6, Appli
30	100.5	3.4	412	15	US-10-156-761-9945	Sequence 9945, Ap
31	100.5	3.4	506	9	US-09-864-761-46512	Sequence 46512, A
32	100.5	3.4	1356	15	US-10-090-280-33	Sequence 33, Appl
33	100.5	3.4	1359	15	US-10-090-280-34	Sequence 34, Appl
34	98.5	3.4	411	9	US-09-828-746-2	Sequence 2, Appli
35	98.5	3.4	411	15	US-10-121-746-83	Sequence 83, Appl
36	98.5	3.4	500	12	US-10-247-813-6	Sequence 6, Appli
37	98.5	3.4	500	12	US-10-247-813-28	Sequence 28, Appl
38	98	3.3	458	15	US-10-001-073-46	Sequence 46, Appl
39	98	3.3	459	10	US-09-738-626-6998	Sequence 6998, Ap
40	98	3.3	461	15	US-10-225-567A-44	Sequence 44, Appl
41	98	3.3	462	15	US-10-001-073-44	Sequence 44, Appl
42	97.5	3.3	338	10	US-09-886-055-205	Sequence 205, App
43	97.5	3.3	338	11	US-09-804-291-205	Sequence 205, App
44	97.5	3.3	338	12	US-10-017-161-30	Sequence 30, Appl
45	97.5	3.3	808	15	US-10-223-047-33	Sequence 33, Appl

ALIGNMENTS

RESULT 1
US-09-973-180-2
; Sequence 2, Application US/09973180
; Publication No. US20030082553A1
; GENERAL INFORMATION:
; APPLICANT: Oestra, Ben Heutink, P Duijin, C.M.
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS FOR AUTOSOMAL DOMINANT HEMOCHROMATOSIS
; FILE REFERENCE: 24584
; CURRENT APPLICATION NUMBER: US/09/973,180
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 60/301,429
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-973-180-2

Query Match	99.8%	Score	2924	DB	11	Length	571
Best Local Similarity	99.8%	Pred	No	3	7e-272		
Matches	570	Conservative	1	Mismatches	0	Indels	0
Gaps	0						
Qy	1	MTRAGDHNRRGCCGSLADYLTSAKFLLYLGHSLSTWGDNRMHFAVSFLVELYGNLSLL	60				
Db	1	MTRAGDHNRRGCCGSLADYLTSAKFLLYLGHSLSTWGDNRMHFAVSFLVELYGNLSLL	60				
Qy	61	TAVYGLVAGSVLVILGAIIGDWDKNARLKVQAOTSLVVQNVSVILCGIILMMVFLKHKL	120				
Db	61	TAVYGLVAGSVLVILGAIIGDWDKNARLKVQAOTSLVVQNVSVILCGIILMMVFLKHKL	120				
Qy	121	LTWYHGVLTSCYLIITITIANIANLASTATAITQTDWIVVWAGEDRSKLANMATTIRI	180				
Db	121	LTWYHGVLTSCYLIITITIANIANLASTATAITQTDWIVVWAGEDRSKLANMATTIRI	180				
Qy	181	DQLTNILAPVAGQITFGSPVIGCGFISGNLNVSMCVYLLMKVYQKTPALAVKAGLK	240				
Db	181	DQLTNILAPVAGQITFGSPVIGCGFISGNLNVSMCVYLLMKVYQKTPALAVKAGLK	240				

Db 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLMKVYQKTPALAVKAGLK 240
Qy 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300
Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300
Qy 301 SYYNQPVFLAGMGLAFLYMTVLGFDCTTGGYAYTQGLSGSLILMGSASAITGIMGTVA 360
Db 301 SYYNQPVFLAGMGLAFLYMTVLGFDCTTGGYAYTQGLSGSLILMGSASAITGIMGTVA 360
Qy 361 TWLRRKCGLVRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPFEDIRSRFIQGESITP 420
Db 361 TWLRRKCGLVRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPFEDIRSRFIQGESITP 420
Qy 361 TWLRRKCGLVRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPFEDIRSRFIQGESITP 420
Db 361 TWLRRKCGLVRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPFEDIRSRFIQGESITP 420
Qy 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLLFAGVIAARIGLWSFDLTVTQLL 480
Db 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLLFAGVIAARIGLWSFDLTVTQLL 480
Qy 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLISVSFVAMGHIMYFR 540
Db 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLISVSFVAMGHIMYFR 540
Qy 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
Db 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
RESULT 2
US-09-715-927-6-3
; Sequence 3, Application US/09973180
; Publication No. US20030082553A1
; GENERAL INFORMATION:
; APPLICANT: Oestra, Ben Heutink, P Duijin, C.M.
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS FOR AUTOSOMAL DOMINANT HEMOCHROMATOSIS
; FILE REFERENCE: 24584
; CURRENT APPLICATION NUMBER: US/09/973,180
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 60/301,429
; PRIOR FILING DATE: 2001-06-29
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: 144
; OTHER INFORMATION:
US-09-715-927-6-3
Query Match 99.8%; Score 2924; DB 11; Length 571;
Best Local Similarity 99.8%; Pred. No. 3.7e-272;
Matches 570; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTRAGDHNRRGCGSLADYLTSAKFLYLGHSLSTWGDWRMHFAVSFVFLVLYGNSLL 60
Db 1 MTRAGDHNRRGCGSLADYLTSAKFLYLGHSLSTWGDWRMHFAVSFVFLVLYGNSLL 60
Qy 61 TAVYGLVAVGSVLVGLGAIIGDWDVKNARLKVQTSLVVQNVSVILCGIILMVFLKHHEL 120
Db 61 TAVYGLVAVGSVLVGLGAIIGDWDVKNARLKVQTSLVVQNVSVILCGIILMVFLKHHEL 120
Qy 121 LTMHGWLTSCYLIITITIANIANLASTATAITIQRDWIVVWAGEDRSKLANNNATIRRI 180
Db 121 LTMHGWLTSCYLIITITIANIANLASTATAITIQRDWIVVWAGEDRSKLANNNATIRRI 180
Qy 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLMKVYQKTPALAVKAGLK 240
Db 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLMKVYQKTPALAVKAGLK 240
Qy 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300
Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300

Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300
Qy 301 SYYNQPVFLAGMGLAFLYMTVLGFDCTTGGYAYTQGLSGSLILMGSASAITGIMGTVA 360
Db 301 SYYNQPVFLAGMGLAFLYMTVLGFDCTTGGYAYTQGLSGSLILMGSASAITGIMGTVA 360
Qy 361 TWLRRKCGLVRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPFEDIRSRFIQGESITP 420
Db 361 TWLRRKCGLVRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPFEDIRSRFIQGESITP 420
Qy 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLLFAGVIAARIGLWSFDLTVTQLL 480
Db 421 TKIPEITTEIYMSNGSNSANIVPETSPEVPIISVSLLFAGVIAARIGLWSFDLTVTQLL 480
Qy 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLISVSFVAMGHIMYFR 540
Db 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLISVSFVAMGHIMYFR 540
Qy 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
Db 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
RESULT 3
US-09-715-927-6-4
; Sequence 4, Application US/09973180
; Publication No. US20030082553A1
; GENERAL INFORMATION:
; APPLICANT: Oestra, Ben Heutink, P Duijin, C.M.
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS FOR AUTOSOMAL DOMINANT HEMOCHROMATOSIS
; FILE REFERENCE: 24584
; CURRENT APPLICATION NUMBER: US/09/973,180
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: 60/301,429
; PRIOR FILING DATE: 2001-06-29
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: 144
; OTHER INFORMATION: Xaa = any amino acid except Asp
US-09-715-927-6-4
Query Match 99.8%; Score 2922; DB 11; Length 571;
Best Local Similarity 99.8%; Pred. No. 5.8e-272;
Matches 570; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MTRAGDHNRRGCGSLADYLTSAKFLYLGHSLSTWGDWRMHFAVSFVFLVLYGNSLL 60
Db 1 MTRAGDHNRRGCGSLADYLTSAKFLYLGHSLSTWGDWRMHFAVSFVFLVLYGNSLL 60
Qy 61 TAVYGLVAVGSVLVGLGAIIGDWDVKNARLKVQTSLVVQNVSVILCGIILMVFLKHHEL 120
Db 61 TAVYGLVAVGSVLVGLGAIIGDWDVKNARLKVQTSLVVQNVSVILCGIILMVFLKHHEL 120
Qy 121 LTMHGWLTSCYLIITITIANIANLASTATAITIQRDWIVVWAGEDRSKLANNNATIRRI 180
Db 121 LTMHGWLTSCYLIITITIANIANLASTATAITIQRDWIVVWAGEDRSKLANNNATIRRI 180
Qy 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLMKVYQKTPALAVKAGLK 240
Db 181 DQTNILAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLMKVYQKTPALAVKAGLK 240
Qy 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300
Db 241 EEETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHEQPTCASQMAEPFRTRDGVW 300
Qy 301 SYYNQPVFLAGMGLAFLYMTVLGFDCTTGGYAYTQGLSGSLILMGSASAITGIMGTVA 360
Db 301 SYYNQPVFLAGMGLAFLYMTVLGFDCTTGGYAYTQGLSGSLILMGSASAITGIMGTVA 360

Db 301 SYNQPVFLAGMGLAFYMTVLGFDCTTGYAYTQGLSGSILSLMGASAITGIMGTAVF 360
 Qy 361 TWLRKCGLVRTGLISGLAQSLCLILCVISVFMPSPLDLSVSPEDIRSRFIQESITP 420
 Db 361 TWLRKCGLVRTGLISGLAQSLCLILCVISVFMPSPLDLSVSPEDIRSRFIQESITP 420
 Qy 421 TKIPITTEIYMSNGSANSANIVPETSPEVPIISVSLIFAGVIAAIGLWSFDLTVTQLL 480
 Db 421 TKIPITTEIYMSNGSANSANIVPETSPEVPIISVSLIFAGVIAAIGLWSFDLTVTQLL 480
 Qy 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLSVSVFAMGHIMYFR 540
 Db 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLSVSVFAMGHIMYFR 540
 Qy 541 FAQNTLGNKLFACGPDADKEVRKENQANTSVV 571
 Db 541 FAQNTLGNKLFACGPDADKEVRKENQANTSVV 571

RESULT 4
 US-09-729-674-130
 ; Sequence 130, Application US/09729674
 ; Patent No. US20010039335A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: LaVallie, Edward R.
 ; APPLICANT: Collins-Racie, Lisa A.
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Merbers, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Agostino, Michael J.
 ; APPLICANT: Steindinger II, Robert J.
 ; APPLICANT: Spaulding, Vikki
 ; APPLICANT: Wong, Gordon G.
 ; APPLICANT: Clark, Hilary
 ; APPLICANT: Fectel, Kim
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 ; FILE REFERENCE: 6055-64X
 ; CURRENT APPLICATION NUMBER: US/09/729,674
 ; CURRENT FILING DATE: 2000-12-04
 ; PRIOR APPLICATION NUMBER: 09/539,330
 ; PRIOR FILING DATE: 2000-03-30
 ; NUMBER OF SEQ ID NOS: 283
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 130
 ; LENGTH: 571
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: (202)
 ; NAME/KEY: UNSURE
 ; LOCATION: (504)
 US-09-729-674-130

Query Match 99.6%; Score 2917; DB 9; Length 571;
 Best Local Similarity 99.6%; Pred. No. 1.7e-271;
 Matches 569; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 MTRAGDHNQRQCCGSLADYLTSAKFLYLGLHSLSTWGDNRWHFAVSFLVELYGNLSLL 60
 Db 1 MTRAGDHNQRQCCGSLADYLTSAKFLYLGLHSLSTWGDNRWHFAVSFLVELYGNLSLL 60
 Qy 61 TAVYGLVAVGSLVLGAILGWDKNAKLQAQTSLVVQNVSVILCGIILMMVFLKHHEL 120
 Db 61 TAVYGLVAVGSLVLGAILGWDKNAKLQAQTSLVVQNVSVILCGIILMMVFLKHHEL 120
 Qy 121 LTMHGWLTSCYILIIITIANIANLASTATAITQDRIWVIVVAGEDRSKLANMATIRRI 180
 Db 121 LTMHGWLTSCYILIIITIANIANLASTATAITQDRIWVIVVAGEDRSKLANMATIRRI 180

Qy 181 DQNLNILAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLK 240
 Db 181 DQNLNILAPMAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLK 240
 Qy 241 BEETELQNLNHLKDEPKPLEGTHLMGVKDSNIHELEHEQEPPTCASQMAEPFRFRDQWV 300
 Db 241 BEETELQNLNHLKDEPKPLEGTHLMGVKDSNIHELEHEQEPPTCASQMAEPFRFRDQWV 300
 Qy 301 SYNQPVFLAGMGLAFYMTVLGFDCTTGYAYTQGLSGSILSLMGASAITGIMGTAVF 360
 Db 301 SYNQPVFLAGMGLAFYMTVLGFDCTTGYAYTQGLSGSILSLMGASAITGIMGTAVF 360
 Qy 361 TWLRKCGLVRTGLISGLAQSLCLILCVISVFMPSPLDLSVSPEDIRSRFIQESITP 420
 Db 361 TWLRKCGLVRTGLISGLAQSLCLILCVISVFMPSPLDLSVSPEDIRSRFIQESITP 420
 Qy 421 TKIPITTEIYMSNGSANSANIVPETSPEVPIISVSLIFAGVIAAIGLWSFDLTVTQLL 480
 Db 421 TKIPITTEIYMSNGSANSANIVPETSPEVPIISVSLIFAGVIAAIGLWSFDLTVTQLL 480
 Qy 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLSVSVFAMGHIMYFR 540
 Db 481 QENVIESERGIINGVQNSMNYLLDLHFMVILAPNPEAFGLLVLSVSVFAMGHIMYFR 540
 Qy 541 FAQNTLGNKLFACGPDADKEVRKENQANTSVV 571
 Db 541 FAQNTLGNKLFACGPDADKEVRKENQANTSVV 571

RESULT 5
 US-09-809-391-460
 ; Sequence 460, Application US/09809391
 ; Publication No. US20030049618A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ruben et al.
 ; TITLE OF INVENTION: 186 Human Secreted proteins
 ; FILE REFERENCE: P2002P2
 ; CURRENT APPLICATION NUMBER: US/09/809,391
 ; CURRENT FILING DATE: 2001-03-16
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 761
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 460
 ; LENGTH: 383
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (383)
 ; OTHER INFORMATION: Xaa equals stop translation
 US-09-809-391-460

Query Match 61.1%; Score 1789; DB 11; Length 383;
 Best Local Similarity 92.4%; Pred. No. 2.8e-163;
 Matches 353; Conservative 3; Mismatches 26; Indels 0; Gaps 0;
 Qy 190 MAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLKEESTELKOL 249
 Db 1 MAVGQIMTFGSPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLKEESTELKOL 60
 Qy 250 NLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPPTCASQMAEPFRFRDQWVSYNQPVFL 309
 Db 61 NLHKDTEPKPLEGTHLMGVKDSNIHELEHEQEPPTCASQMAEPFRFRDQWVSYNQPVFL 120
 Qy 310 AGMGLAFYMTVLGFDCTTGYAYTQGLSGSILSLMGASAITGIMGTAVFTWLRRKCGL 369
 Db 121 AGMGLAFYMTVLGFDCTTGYAYTQGLSGFHPQYFDGSIYNNNGNSCYLATSQWVF 180
 Qy 370 VRTGLISGLAQSLCLILCVISVFMPSPLDLSVSPEDIRSRFIQESITPTKIPETITE 429
 Db 181 GSAGLISGLAQSLCLILCVISVFMPSPLDLSVSPEDIRSRFIQESITPTKIPETITE 240
 Qy 430 IYMSNGSANSANIVPETSPEVPIISVSLIFAGVIAAIGLWSFDLTVTQLLQENVIESER 489

Db 241 IYNSGNSANIWPETSPESVPIISVLSLLPAGVIARIGLWSFDLTVTQLQENVIESR 300
QY 490 GIINGVQNSMYLLDLLHFMVILAPNPEAFGLLVLSVSVFVAMGHIMYFRFAQNTLGNK 549
Db 301 GIINGVQNSMYLLDLLHFMVILAPNPEAFGLLVLSVSVFVAMGHIMYFRFAQNTLGNK 360
QY 550 LFACGPDPAKEVRKENQANTSVV 571
Db 361 LFACGPDPAKEVRKENQANTSVV 382

RESULT 6
US-09-882-171-460
; Sequence 460, Application US/09882171
; Publication No. US20030175859A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P2
; CURRENT APPLICATION NUMBER: US/09/882,171
; CURRENT FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 09/809,391
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 09/149,476
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/04493
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: 60/040,162
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,333
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/038,621
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,626
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,334
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,336
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/040,163
; PRIOR FILING DATE: 1997-03-07
; PRIOR APPLICATION NUMBER: 60/047,600
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,615
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,597
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,502
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,633
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,583
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,617
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,618
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,503
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,592
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,581
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,584
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,500
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,587
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,492
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,598

; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,613
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,582
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,596
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,612
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,632
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/047,601
; PRIOR FILING DATE: 1997-05-23
; PRIOR APPLICATION NUMBER: 60/043,580
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,568
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,314
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,569
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,311
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,671
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,674
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,669
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,312
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,313
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,672
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/043,315
; PRIOR FILING DATE: 1997-04-11
; PRIOR APPLICATION NUMBER: 60/048,974
; PRIOR FILING DATE: 1997-06-06
; PRIOR APPLICATION NUMBER: 60/056,886
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,877
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,889
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,893
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,630
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,878
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,662
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,872
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,882
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,637
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,903
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,888
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,879
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,880
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,894
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,911
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 60/056,636
; PRIOR FILING DATE: 1997-08-22

Query Match	61.1%;	Score 1789;	DB 12;	Length 383;
Best Local Similarity	92.4%;	Pred. No. 2.8e-163;		
Matches 353; Conservative	3;	Mismatches 26;	Indels 0	

RESULT 7

Query Match 4.3%; Score 125.5; DB 15; Length 575;
Best Local Similarity 20.4%; Pred. NO. 0.0049;
Matches 102; Conservative 70; Mismatches 162; Indels 165; Gaps 23;

39	QY	DRMHFAVSFVFLVELYGNLSLLTA-----VYGLVVAGSVLVLGAIIGDWDVKNA	87
57	Db	DLQW--VIDGYTPVAAALMLLGAIGDKYSRRGALVLGLVVFQG---GAVAGSLVDSAT	110
88	QY	RLKVAQTSLVVQVNSVILCGIILMMVFLKHKE-----LLTM	123
111	Db	GVTAARVMGVGAALIMPATLSLAATFPPRAERAKAITLTATAGLAAAGPLVAGALLQ	170
124	QY	YHGCVLTSCVILLIITIANIANLASTATAITIQRDWIVVAGEDRSKLANNATIRRIDQL	183
171	Db	NHGW--SSTFLLINPFAFAVIGT-----FVLVPPSRA-----AHHRIDYV	210
184	QY	TNIIAPMAVGQI--MTFGSEVIGGETISGNLVSMCMVEYVLL--WKVYQKTPALAVKAGL	239

Db 211 GGLSVVWIGSLVYMIIEGPHFGVGVKAVTAATAAAGVGVFLFGWELRHRPVL----- 264
QY 240 KEBETELKQLNLHKDTEPKPLEGTHLMGVKDSNIHELEHQETPCASQMAEPRTFRDGW 299
Db 265 -----DVRFAHRR--FAGSNLA----- 280
QY 300 VSYNQPVFLAGGLAFLYMT-----VLGFDCITTTGAYVYQGLSGSLTSLMGASALTGI 354
Db 281 VALF-----FLAVFG-ABYYLUTQHLQFVLGVYDALDTGVRMLP-LAGAVFV-----GSALTG- 329
QY 355 MGTVAFTLRRKCGL-----VRLTGISGLAQLSCILC-----VISVFMPSGPLDSL 402
Db 330 -----YLTPRVGMKITVAGVMGGTAALALLTRVDAASSYGGDFVLPLVILGLAIGLAL 382
QY 403 SPED-IRSPFOGESITPKIPEITTEIYNS-----NGSNSANIVPTSPESVP 451
Db 383 SPCTDAIMGAPPEALGVGAVNDTSLGGLGIALILGSLVLAGSYSSHLSDATTSKLP 442
QY 452 IISVS-----LLFAGVIAAR 466
Db 443 ASALQAQDSVVGAGYAVAR 461

RESULT 8

US-10-283-423-6
; Sequence 6, Application US/10283423.
; Publication No. US20030162223A1
; GENERAL INFORMATION:
; APPLICANT: Lowery, David E.
; APPLICANT: Smith, Valdin G.
; APPLICANT: Kubiak, Teresa M.
; APPLICANT: Larsen, Martha J.
; TITLE OF INVENTION: Drosophila G Protein Coupled Receptors, Nucleic Acids, And Method
; TITLE OF INVENTION: Related To The Same
; FILE REFERENCE: PHRM0002-102
; Application Project
; -----
; CURRENT APPLICATION NUMBER: US/10/283,423
; CURRENT FILING DATE: 2002-10-30
; Earlier Applications
; -----
; PRIOR APPLICATION NUMBER: PriorAppNumber : 09/693,746
; PRIOR FILING DATE: PriorFilingDate : 2000-10-20
; PRIOR APPLICATION NUMBER: PriorAppNumber : 09/425,676
; PRIOR FILING DATE: PriorFilingDate : 1999-10-22
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 458
; TYPE: PRT
; ORGANISM: D. melanogaster
US-10-283-423-6

Query Match 4.1%; Score 120; DB 12; Length 458;
Best Local Similarity 20.1%; Pred. No. 0.012;
Matches 77; Conservative 60; Mismatches 98; Indels 148; Gaps 21;

QY 5 GDHNRQCGCGSLADYLTSAKFLLYLG-----HSLSTWG----- 38
Db 12 GEH-----LSGYASSNSVRYLDDRHPLDYLDLGTVHALNTTAINSTDNETGSRP 62
QY 39 -----DRMWHFAVSVELVELYGNLSLLTAVYGLVAVAGSVLVGLAIGDWW 83
Db 63 LDPVLIDRFUSNRAVDSPWYH-----MLISMYG-----VLIVFGAL--GNTLWVIAVIRKPI 112
QY 84 DKNARLKVACTSLVQVNVSV---ILCGIILMMVFLHKHELLTMYHGWVLTSCYILIITIA 140
Db 113 MPTAR-----NLFILNLAISSDLLCLVTPMLTLM---EILSKY--WPYGSCSILCKTIA 161
QY 141 NIANL-----ASTATAITIQDWMVIVVAGEDRSKLANMNATIRRIDQLTNILA-PMAV-G 193
Db 162 MLQALCIVFVSTISITAIAFDQYVIVVPTSDLSLQFVGAVTILAGIWWALLASPLFVYK 221

QY 194 QIMTGGSPV---IG-----CGRISGW-----NLVSMCVVEY-----VLLWK 225
Db 222 EINTDTPALLOQIGLQDTIPYC--IEDWPSRNGRFYISFSLCQVLYLPILIVSVAYFG 279
QY 226 VYQKTPA-----LAVKAGLKEETE-----LKQNLHKDTE 256
Db 280 IYNKLKSRITTVAVQASSAQKVERGRMRKRTNCLLSIAITIFGVSWLPLNFFNLVADME 339
QY 257 PKPLEGT-----HLMGVKDS 271
Db 340 RSPVTQSMVLRYAICHMIGMSSA 362

RESULT 9

US-10-213-821-6
; Sequence 6, Application US/10213821
; Publication No. US20030180297A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; APPLICANT: Lowery, David E.
; APPLICANT: Smith, Valdin G.
; APPLICANT: Kubiak, Teresa M.
; APPLICANT: Larsen, Martha J.
; TITLE OF INVENTION: Drosophila G Protein Coupled Receptors, Nucleic Acids, And Methods
; TITLE OF INVENTION: Related To The Same
; FILE REFERENCE: 6297.2Cp
; CURRENT APPLICATION NUMBER: US/10/213,821
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: 09/693,746
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/425,676
; PRIOR FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 458
; TYPE: PRT
; ORGANISM: D. melanogaster
US-10-213-821-6

Query Match 4.1%; Score 120; DB 12; Length 458;
Best Local Similarity 20.1%; Pred. No. 0.012;
Matches 77; Conservative 60; Mismatches 98; Indels 148; Gaps 21;

QY 5 GDHNRQCGCGSLADYLTSAKFLLYLG-----HSLSTWG----- 38
Db 12 GEH-----LSGYASSNSVRYLDDRHPLDYLDLGTVHALNTTAINSTDNETGSRP 62
QY 39 -----DRMWHFAVSVELVELYGNLSLLTAVYGLVAVAGSVLVGLAIGDWW 83
Db 63 LDPVLIDRFUSNRAVDSPWYH-----MLISMYG-----VLIVFGAL--GNTLWVIAVIRKPI 112
QY 84 DKNARLKVACTSLVQVNVSV---ILCGIILMMVFLHKHELLTMYHGWVLTSCYILIITIA 140
Db 113 MPTAR-----NLFILNLAISSDLLCLVTPMLTLM---EILSKY--WPYGSCSILCKTIA 161
QY 141 NIANL-----ASTATAITIQDWMVIVVAGEDRSKLANMNATIRRIDQLTNILA-PMAV-G 193
Db 162 MLQALCIVFVSTISITAIAFDQYVIVVPTSDLSLQFVGAVTILAGIWWALLASPLFVYK 221
QY 194 QIMTGGSPV---IG-----CGRISGW-----NLVSMCVVEY-----VLLWK 225
Db 222 EINTDTPALLOQIGLQDTIPYC--IEDWPSRNGRFYISFSLCQVLYLPILIVSVAYFG 279
QY 226 VYQKTPA-----LAVKAGLKEETE-----LKQNLHKDTE 256
Db 280 IYNKLKSRITTVAVQASSAQKVERGRMRKRTNCLLSIAITIFGVSWLPLNFFNLVADME 339
QY 257 PKPLEGT-----HLMGVKDS 271
Db 340 RSPVTQSMVLRYAICHMIGMSSA 362


```

Qy 457 ----LLFAGVIAAIRGLWSFDLTVTQLQENVIESERGIINGVQ 496
      :      |||      :      :      :      :      :      :
Db 453 AGDMVFIPDVIEVPAGK--SLEVTMLNEDDMVHDLK-FANGVQ 497
      :      |||      :      :      :      :      :      :

RESULT 11
US-09-738-626-3641
; Sequence 3641, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: Patent in ver. 3.0
; SEQ ID NO 3641
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3641

```

Query Match	3.9%;	Score 114.5;	DB 10;	Length 424;
Best Local Similarity	21.4%;	Pred. No. 0.035;		
Matches	86;	Conservative	55;	Mismatches 131; Indels 129; Gaps 16;

Qy	28	LYLGHSLSWTGDRMHPFAVSFLVELYGNs--LLLTAyVGLVAGSVLVLGAIIGDWVK	85
Db	21	LWLGNVATVGAQLTVVAVPVQYQMTGSSGYVGLTGLGLPIVIFGLYGGSTADAPDK	80
Qy	86	NARLKVAQTSLVVQNVSVISCGIILMMVFLHKHELLTMYHGWLt-----SCYLIITi-	139
Db	81	RI-----VLCTIGMCV-----TTAGFWLTLILGNEINWILLNFS	117
Qy	140	---ANTIANLASTATITQRDWIVVWAGEDRSKLANNATIRRIDQLTNIIAPMAVGIM	196
Db	118	LQQAFFAVNPQTRAIL--RSILPI-----DQLASATSLNMLLMQTAIVGPIIAGALI	169
Qy	197	TFGSPVTCGCFISGWNILVSMCEVULLWKVYQ----KTPALAVKAGLKEEETELKQLNIH	252
Db	170	-----PLIGFWLYFLDVSNI---IPTLWAVNSUPSIRKPSGKVMKAG-----	208
Qy	253	KDTEPKPLEGTHLMGVKSDSIHLEHQBECTCASQMAEPFTRFDGWMVSYYNQPVFLAGM	312
Db	209	-----FASVVDGLKYLGAQPVLL---	226
Qy	313	GLAFLYVWVLGFPDCTTGAYATQGLSGSIILMGASAITGIMGTVAFTHLRRKGLVRT	372
Db	227	-----MNVLDLIAMIFGMPRALYPEIAEVNFVGGDGAGATM--LAFMYSSMAVGAVLG	277
Qy	373	GLISG-LAQLS-----CLILCVISVFMPGSPFLDLSVSP	404
Db	278	GVLSUGWVARISROGVAVYWCIIANGAAVALGG--VAIVVSP	316

RESULT	12
--------	----

```

US-09-769-744A-120
; Sequence 120, Application US/09769744A
; Publication No. US20030134407A1
; GENERAL INFORMATION:
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hanniffy, Sean B
; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21122WO
; CURRENT APPLICATION NUMBER: US/09/769, 744A
; PRIOR FILING DATE: 2001-01-26
; PRIOR FILING DATE: PCT/GB99/02452
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: GB 9816336.3
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125329
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 120
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-744A-120

Query Match      3.9%; Score 114; DB 12; Length 392;
Best Local Similarity 20.1%; Pred. No. 0.035;
Matches 77; Conservative 67; Mismatches 132; Indels 108; Gaps 17;

QY 30 LGHSLSTGDRMWHFAVSVF---LVELYGNLSLLTAVYGLVAGSVLV--LGAIGDWD 84
Db 17 MGNLDYGNVWLASMGTTGQVIGYQISELV-----SLVNPFGVIS---D 64

QY 85 KNARLKVAQTSLVVQNVSVILCGIILMMVFLHKLHLLTMVHGWLTSYLIITIANIAN 144
Db 65 RFSRKLMTADLV-----CGILCLAFIRND-----SNWIGA-----LIVANIVQ 106

QY 145 LASTATAITIQDWMVIVVAGEDSKLANMNATIRIDQLTNLA-PMVAV-GQIM 204
Db 107 AIAPAFSRTANKAITIEVEKD--EIVYINSRLSELVLQVVGVSPLVFLVLOFAS--- 160

QY 205 CGFTSGNVLVSMCVYVLLMKVQKTPALAVKAGLKEEETELKQLNLHDKTEPKPLEGTH 264
Db 161 -----LHMTLLDLSLTFFAVIVAFLPKEAKVQ-----EKKATGRD 199

QY 265 L-MGVKDSNIHELEHQE-----PTCASQMAEPFRTPRDGVSVYVQVFLAGMGLAFLYM 319
Db 200 IFVDIKD-GLHYIWHQOEIFFLLLVASSVNVFFFAAF-EFLLPFSNQL-----YGSEGAYA 252

QY 320 TVLGFDCITTYGAYTQGLSGSILSILMGASAITGIMGTVAFTWLRKCGLVRTGLISGLA 379
Db 253 SIL-----TMAIGSITGALLASKIKANIVL----- 279

QY 380 QLSCLILCVISVFMGSPDLDSVS 403
Db 280 -LILLALTGVGVFMGSLPLTFLS 302

RESULT 13
US-09-804-551B-38
; Sequence 38, Application US/09804551B
; Patent No. US20020056151A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Receptors for peptides from insects
; FILE REFERENCE: Le A 34 394
; CURRENT APPLICATION NUMBER: US/09/804, 551B
; CURRENT FILING DATE: 2001-03-12
; PRIOR FILING DATE: 2000-03-18
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 38
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-804-551B-38

Query Match      3.9%; Score 113; DB 9; Length 522;
Best Local Similarity 20.0%; Pred. No. 0.067;
Matches 76; Conservative 58; Mismatches 98; Indels 148; Gaps 20;

QY 5 GDHNRQRCGCGSLADYLTSAKFLLYLG-----HSLSTWG----- 38
Db 59 GEH-----LSGVASSNSVRYLDRHPLDYLDTGVTVALNTTAINTSDLNETGSRP 109

QY 39 -----DRMWHFAVSFLVELYGNLSLLTAVYGLVAGSVLVILGAIIGDVM 83
Db 110 LDPVLIDRELSNRAVDSPWYH-----MLTSMYG-----VLIVFGAL--GNTLVVIAVIRKPI 159

QY 84 DKNARLKVAQTSLVVQNVSVILCGIILMMVFLHKLHLLTMVHGWLTSYLIITIANIA 143
Db 160 METAR-----NLFINLAISVTMPLTM-----EILSKY--WPYGSCSILCKTIAMIQ 205

QY 144 NL-----ASTATAITIQDWMVIVVAGEDSKLANMNATIRIDQLTNLA-PMVAV-GQIM 196
Db 206 ALCIFVSTISITAITAFDRYQVIVVPTDLSQFVGAVTILAGIWAALLLASPLFVYKELI 265

QY 197 TFGSPV---IG-----CGFISGW-----NLVSMCVY-----VLLWKVYQ 228
Db 266 NDTPTALLQOIGLQDTPYPC--IEDWPSRNGRFFYSIFSLCVQYLVLPILIVSVAYEGINV 323

QY 229 KTPA---LAVKAGLKEETE-----HLMGVKDS 271
Db 324 KUKSRITVVAVQSSAQKVERGRMKRTNCLLSIAIIFGVSWLPFLNPNLYADMERSP 383

QY 260 LEGT-----HLMGVKDS 271
Db 384 VTQSMVLRVYAICHMIGMSSA 403

RESULT 14
US-10-419-190-2
; Sequence 2, Application US/10419190
; Publication No. US20030166155A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; FILE REFERENCE: AND USES THEREOF
; FILE REFERENCE: CL000872CON
; CURRENT APPLICATION NUMBER: US/10/419,190
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: 60/232,856
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/685,852
; PRIOR FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Human
US-10-419-190-2

Query Match      3.7%; Score 108.5; DB 12; Length 486;
Best Local Similarity 22.3%; Pred. No. 0.16;
Matches 106; Conservative 65; Mismatches 176; Indels 129; Gaps 25;

QY 46 VSVFLVEL---YGNLSLLTA-VYGLVAGSVLVILGAIIGDWDVKNARKV-----AQ 93
Db 39 ISIFFVEFTQTYTQDYAQTAWIHSIV--DCVTMLCAPGSGVSVNHLSCQVGMGLGLLAS 96
QY 94 TSLVVQNVSVILCGIILMMVFLHKLHLLTMVHGWLTS-----CYLIITIANIANLAST 148

```

Db 97 TGLILSFATSL-----KHLXLT-----GVLTLGLFALCY--SPAIAVMGKYFSR 140
Qy 149 ATAITIQDWIIVVAGEDRSKLANNNATIRIDQLTNIIAP---MAVGQIMTFGSPVIGC 205
Db 141 RKALA-----YGIAMSGSIG-----TFILAPVQVLLIQFWSRGALLILG 181
Qy 206 GFISGNLVMCEVYVLLKWKYKTPALAVKAGLKEETEELKQLNLHKDTPEKPLEGTHL 265
Db 182 GFV-----LNLGV-----CGALMRPITLKEDHTTPEQNHVCR-TQKEDIK---- 220
Qy 266 MGVKDSNIHELEHEOEP---CASQMAEPRTFRDGMVSYNQPVFLA-GMGLAFLYMTV 321
Db 221 ---RVSPSSLTKEWAQCLCCLQOEYFELMSD-FVLAVSVLFMAYGCSPLFVYL-- 274
Qy 322 LGFDCITTYAYTOGLSGSILSILMGASAITGIMGTVAFTWLRK-----CGLVRT 372
Db 275 -----VPYALSVGVSHQQAFLMSILGVIDIIGNITFGWLTDRCLKNYQVVCYLEAV 327
Qy 373 GLISGLAQLSCLILCVISVFMFGSPDLDSVPFFEDIRSRFIOGESITPTKIPETITEYM 432
Db 328 GM-DGLCYLCPLMQLSLPLVPFS-----CTFGYFDGAYVTLIPVVTTEIVG 373
Qy 433 SNGNSA-----NIVPE-TSPESVPIISVSLFAGVIAARIGLWSFDLTVTQLL 480
Db 374 TTSLSALGVVFLHAVPYLVSP---PIAGRLVDTGTSYTAFLLCGFSMIFSSVL 426

RESULT 15

US-09-995-007-2
; Sequence 2, Application US/09995007
; Patent No. US20020123102A1
; GENERAL INFORMATION:
; APPLICANT: NANCY CARRASCO, ET AL.
; TITLE OF INVENTION: THYROID SODIUM/IODIDE SYMPORTER AND
; NUCLEIC ACID ENCODING SAME
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: AMSTER, ROTHSTEIN & EBENSTEIN
; STREET: 90 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH 1.44 Mb STORAGE
; DISKETTE
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/995,007
; FILING DATE: 26-NO. US20020123102A1-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/595,553A
; FILING DATE: FEBRUARY 1, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAIG J. ARNOLD
; REGISTRATION NUMBER: 34,287
; REFERENCE/DOCKET NUMBER: 96700/393
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-5995
; TELEFAX: (212) 286-0854 or 286-0082
; TELE: TWX 710-581-4766
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 618
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: <Unknown>
; DESCRIPTION: PROTEIN
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: RAT

INDIVIDUAL ISOLATE: SODIUM/IODIDE SYMPORTER
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-995-007-2

Query Match 3.7%; Score 108; DB 10; Length 618;
Best Local Similarity 19.1%; Pred. No. 0.26;
Matches 103; Conservative 71; Mismatches 160; Indels 204; Gaps 27;

Qy 14 CGSLADYLTSAKPLLYLGHSLSTWGRDMHFAVSFVELVYGNLSLLLTAVVGLVAVGVL 73
Db 132 CGTL-QYLVAT--MLYTG-----IVYAPALILNQVTLGLDIWASLL 169
Qy 74 VLGAIIIGDWDMKARLKVQATSLVQNVSVILCGIILMMVFLHKHELLTMYVHGW-LTSC 132
Db 170 STGIICITLYTTVGKMAVWTD--VFQVVMVLVGF-----WVILARG 209
Qy 133 YILITITANTANLASTATAITIQ-----RDWIVVWAG----- 164
Db 210 VILGSPRNVLSLAQNHSRINLMDFPDPRSRYTFWTFIVGGTLVWLWSMYGVNQAVQVRY 269
Qy 165 ----EDRSKLANNNATIRRIDQLTNIIAPMAVGQIMTFGSPVIGCGFISGMNLSVMCME 219
Db 270 VACHTEGAKALL-----VNQL-----GLFLIVASAAC-CG----- 300
Qy 220 YVLLWKVYQK--TPALAVKAGLKEETEELKQLNLHKDTPEKP-----LEGTH 264
Db 301 --IVMFVYKDCDPLLTGRISAPDOYMPLLVLDIFEDLPVPGFLACAYSGTILSTASTS 358
Qy 265 LMGVKDSNIHELEHEOEP---CASQMAEPRTFRDGMVSYNQPVFLAGMGLAFLYMTVIGF 324
Db 359 INAMAATVEDLIKPRMPLAPR-----KLVFIS-KGLSFIY---GS 396
Qy 325 DCITTGAYTQGLSGSILSILMGASAITG---IMGTV-----APT--WLRKCGL--VR 371
Db 397 ACLTV-----AALSSLLGGVLOGSFTVMGVISGPLLAGTGLMPLPACNTPGVL 446
Qy 372 TGLISGLAQLSCLILCVISVFMFGSPDLDSVPFFEDIRSRFIOGESITPTKIPETITEY 431
Db 447 SGLAAGLA-VSLMVAVGATLYPPGE-----QTMGVLP TSAAC TND SV 488
Qy 432 M---SNGNSANIVPETSPE-----SVPIISVSLFAGVIAARIGLWSFDLTVTQLL 480
Db 489 LLGPPGATNASGNTGIPSSGMDTGRPALADTFYATSYLYGALG-----TUTTML 536

Search completed: November 5, 2003, 18:38:42
Job time : 47 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 18:29:46 ; Search time 29 Seconds
(without alignments)
1893.528 Million cell updates/sec

Title: US-09-715-927-6
Perfect score: 2929
Sequence: 1 MTRAGDHNRRQGCCGLADY.....ACGPDAKEVRKENQANTSVV 571
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	612.5	20.9	524	T02499	hypothetical prote
2	564	19.3	498	T48385	transporter like p
3	503.5	17.4	551	B88949	protein R09B5.4 [i
4	209.5	7.2	677	C89010	protein R08F11.6 [
5	187.5	6.4	482	T01762	hypothetical prote
6	131.5	4.5	523	S06920	glucose transport
7	123.5	4.2	409	C96956	permease [imported
8	123.5	4.2	524	A31318	glucose transporte
9	120.5	4.1	763	A31443	Na+/H+ antiporter
10	118.5	4.0	534	S14144	hexose transport p
11	118	4.0	437	A82140	hypothetical prote
12	117.5	4.0	783	E86254	hypothetical prote
13	116	4.0	469	C90446	permease, multidru
14	115.5	3.9	418	E83686	hypothetical prote
15	115	3.9	503	T26524	hypothetical prote
16	114	3.9	392	B95129	transporter, proba
17	114	3.9	392	G97999	hypothetical prote
18	113.5	3.9	470	A81804	transmembrane effl
19	113.5	3.9	482	B59803	metabolite transpo
20	113.5	3.9	522	A31556	glucose transport
21	113	3.9	435	A2079	hypothetical prote
22	112.5	3.8	435	S65982	transport protein
23	112	3.8	304	B83032	probable permease
24	112	3.8	461	D70073	metabolite transpo
25	112	3.8	1148	H90175	NADH dehydrogenase
26	111.5	3.8	403	E84063	multidrug resistan
27	111.5	3.8	473	G69789	sugar transporter
28	111.5	3.8	679	D69329	conserved hypothet
29	110	3.8	413	A87305	Na+/H+ antiporter

ALIGNMENTS

RESULT 1

T02499

hypothetical protein At2g38460 [imported] - Arabidopsis thaliana

N;Alternate names: hypothetical protein T19C21.5

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 05-Mar-1999 #sequence_revision 05-Mar-1999 #text_change 16-Feb-2001

C;Accession: T02499; C84805

R;Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.M.; Kaul

submitted to the EMBL Data Library, August 1998

A;Description: Arabidopsis thaliana chromosome II BAC T19C21 genomic sequence.

A;Reference number: Z14676

A;Accession: T02499

A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-524 <ROU>

A;Cross-references: EMBL:AC004683; NID:g3395421; PID:g3395426

A;Experimental source: cultivar Columbia

R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; vanKen, S.E.; Unayan, L.; Tallon, L.;

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.B.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A;Reference number: A84420; MUID:20083487; PMID:10617197

A;Accession: C84805

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-524 <STO>

A;Cross-references: GB:AE002093; NID:g3395426; PIDN:AAC28758.1; GSPDB:GN00139

C;Genetics:

A;Gene: At2g38460; T19C21.5

A;Map position: 2

A;Introns: 53/2; 169/2; 330/2; 387/3; 433/3; 470/3

Query Match	20.9%	Score	612.5	DB	2	Length	524
Best Local Similarity	29.1%	Pred. No.	6e-40				
Matches	155	Conservative	95	Mismatches	179	Indels	103
Gaps	9						

Qy	28	LYGHSLSWTGDRMWHFVAVVELVYNSLLTAVYGLVAVGVLVGLAIGDWVDKNA	87
Db	41	LYVGYFLARWSARTWEFSVALYMIHLWPNLSLLAAIYGAIESGSTAIFGPIVGQWVEGMD	100
Qy	88	RLKVAQOTSLAVQNVSVILCG---IILMMVFLHKHELLTMVHGWLTSYIIILITANTAN	144
Db	101	YVKVRLRWLLFQNLSTYIAGGAVIKLLVSLDLSRNLFPV-----FAILVLITNLG	151
Qy	145	----LASTATAITQRDWVWVA-GEORSKIANNATIRRIDLTNLAPMAVQIMTFG	199
Db	152	AIGVSLTLAGTILIERDWAVMVMEGHPPAVLTWNVSIRGIDISSKLLSPVITGLIISFV	211
Qy	200	SPVIGCGFISGNWLVSMCVFVLLWKVYQKTPALAVKAGLKBETELKQLNLHKDTPKP	259
Db	212	SLKASAITFAAWATTATWVEYWLFIYSYGVPAIT-----RSNERRILRSRTKQVEGRDAP	267

Db 454 ITHIMQVITPERBNTVFGMHNALCQTFSLVKDLVILPLPATFAICIFISYGFVSCGH 513
QY 536 IMYFRF-----AONTLGNKLFACGPDKEVRKENQ 565
Db 514 MFPIYLVKTNLSIVGRKLSQL-PE-KKEEKEKE 546

RESULT 4
C89010
protein R08F11.6 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: C89010
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: C89010
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-677 <STO>
A:Cross-references: GB:chr_V; PIDN:AAB54248.1; PID:G2088831; GSPDB:GN00023; CESP:R08F11.
A:Note: similar to mariner transposase
C:Genetics:
A:Gene: R08F11.6
A:Map position: 5

Query Match 7.2%; Score 209.5; DB 2; Length 677;
Best Local Similarity 26.1%; Pred. No. 2.5e-08;
Matches 65; Conservative 47; Mismatches 72; Indels 65; Gaps 11;

QY 135 LIITIANIANLASTATAITIQRDWIVVAGEDRSKLANNNATIRIDQLTNILAPMAVGQ 194
Db 481 LAILFGAINRFLNAEKSMIARDWVVALNG--KNLSRQNAWLTGFDQLNLISPIVVG 538
QY 195 IMT-FG--SPVIGCGFISGNVLNMCVEVLLMKVYQKTPALAVKAGLKEETELKQLNL 251
Db 539 LLTFEGMHQTLIVFGI---YMSLSFLSKLIFLSLT- 571
QY 252 HKDTPKPLEGTHLMGVKDSNIHELEHEQEPTCASQMAEPRTFRD---GNVS-YYNQ 306
Db 572 -----NQNV-SLQNRQEGYCSLPISDEEPLFKPKRPGVISTYWKQT 612
QY 307 VFLAGMGLAFLYMTVLGDCITTYGAYTQGLSGSLSL-NGA-----SAITGIMGTVAFT 361
Db 613 TFCAAFGIALFYKTVMGFDNLAVGYA---TSASDLSVITIGALKSYGAVAGVGVISYA 668
QY 362 WLRRKCGLV 370
Db 669 FLEKRIHII 677

RESULT 5
T01762
hypochemical protein A_IG002P16.6 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 22-Oct-1999
C:Accession: T01762
R:Miller, N.; Beck, C.; Kramer, J.
submitted to the EMBL Data Library, June 1997
A:Reference number: Z14421
A:Reference number: Z14421
A:Accession: T01762
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-482 <MIL>
A:Cross-references: ENBL:AF007270; NID:g2191157; PID:g2191161; GSPDB:GN00063; ATSP:A_IG0
C:Genetics:
A:Gene: ATSP:A_IG002P16.6
A:Map position: 5

A:Introns: 15/2; 86/3; 108/1; 179/3; 238/3; 263/3; 301/1; 396/2; 433/3

Query Match 6.4%; Score 187.5; DB 2; Length 482;
Best Local Similarity 20.2%; Pred. No. 8.5e-07;
Matches 84; Conservative 71; Mismatches 157; Indels 103; Gaps 14;

QY 26 FLLYLGHSLSTWGRDMHFAVSFLVELYGNLSLLTAVYGLVAGSVLVLGAIIGDWDVK 85
Db 107 YVLYASCLVGNLVEQLMNFAPSAIMLY-PSLLPVAVMGFVTKLAIAGGPPVVGKFM 165

QY 86 NARLKVAQTSVLVQNVSVIL-CGIIIMVFLKHELLTMYHGWLVTSCYILLIITIAN-IA 143
Db 166 SPRVPTYISLNVIOAAQVLSAGMI-----THAYTVPSTSSASSILLQPFALLPAGA 220

QY 144 NLASTATAITIQRDWIVVAGEDRS-KLANNNATIRIDQLTNILAPMAVGQIMTFGSPV 202
Db 221 SLGFIASGVALTERDWWLLAGINRPIALAAQNAVLRHLDL----- 261

QY 203 IGCGFISGNVLNMCVEVLLMKVYQKTPAL--AVKAGLKEEETELKQLNLHKDTEPKPL 260
Db 262 -----CETALIWLTKNKFSSGVLDPRPKCSLNSCSAEGSRTN---TDSIYI 302

QY 261 EGTMLMGVKSNIHELEHEQEPTCASQMAEPRTFRDQWVSYNQPVFLAGMGLAFLYMT 320
Db 303 TKYHLL-----CSSVDIGMETIKLGWKEYIQQVLPASLAYVLLYFN 344

QY 321 VLGFDCTITTYGAYTQGLS-----GSIL-----SIIM-----GASA 350
Db 345 IV-----LTPGSLMTAFLTRQCMLLRLGLPLLFSSAWMKIVINLARVFLPITLSNYGLCA 400

QY 351 ITGIMGTVAFTWLRKRCGLVRLTGLISGLAQLSCLILCVISVFMPSGLDLSVSPF 405
Db 401 VMGVAATFSLSANLVKRVGILKVKCFPAICYLE--LAGAVGLFFQASLLAVAVAVY 453

RESULT 6
S06920
glucose transport protein, hepatic - mouse
N:Alternate names: GLUT 2
C:Species: Mus musculus (house mouse)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 24-Sep-1999
C:Accession: S06920; S05319; B44887; 148367
R:Suzue, K.; Lodish, H.F.; Thorens, B.
Nucleic Acids Res. 17, 10099, 1989
A:Title: Sequence of the mouse liver glucose transporter.
A:Reference number: S06920; MUID:90098776; PMID:2602116
A:Accession: S06920
A:Molecule type: mRNA
A:Residues: 1-523 <SUZ>
A:Cross-references: EMBL:X16986; NID:g55511; PIDN:CAA34855.1; PID:g55512
R:Asano, T.; Shibasaki, Y.; Lin, J.L.; Akanuma, Y.; Takaku, F.; Oka, Y.
Nucleic Acids Res. 17, 6386, 1989
A:Title: The nucleotide sequence of cDNA for a mouse liver-type glucose transporter prot
A:Reference number: S05319; MUID:89366666; PMID:2771649
A:Accession: S05319
A:Molecule type: mRNA
A:Residues: 1-105, 'D', 107-180, 'T', 182-326, 'T', 328-350, 'F', 352-431, 'P', 433-523 <ASA>
A:Cross-references: EMBL:X15684; NID:g51090; PIDN:CAA33719.1; PID:g51091
R:Hogan, A.; Heyner, S.; Charron, M.J.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Th
Development 113, 363-372, 1991
A:Title: Glucose transporter gene expression in early mouse embryos.
A:Reference number: A44887; MUID:92111400; PMID:1765007
A:Accession: B44887
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 384-431, 'P', 433-496 <HOG>
A:Cross-references: GB:S77926; NID:g242129; PIDN:AAB20847.1; PID:g242130
A:Experimental source: embryo
A:Note: sequence extracted from NCBI backbone (NCBIN:77926, NCBI:77927)
R:Waeber, G.; Thompson, N.; Haefliger, J.A.; Nicod, P.
J. Biol. Chem. 269, 26912-26919, 1994
A:Title: Characterization of the murine high Km glucose transporter GLUT2 gene and its t
A:Reference number: A55078; MUID:95014557; PMID:7929431

Db 214 NYDLWHILLGSGVRAILQSLLLFFCPSPRYLYIKL-----DEEVKAKQ-SLKR----- 262
Qy 261 EGTILMGVKD--SNIHELEHEOEPTCASQMAEPRTFRDGMVSYNQPVFLAGM-GLAFL 317
Db 263 ----LRGVDDVTKQINERKEREASSQKVIITQLFTN---SSYRQDILVALMLHVAQQ 315
Qy 318 YMTVLGFCITTYAYTQGLSGSILSLMGASAITGIMGTVAFTWLRKRCGLVTRTGLISG 377
Db 316 FSGINGIFVYSTIFQTAGISKPVYATI-GVGAVNMVF-TAVSVFLVEKAGRRSLFLIGM 373
Qy 378 LAQLSCLI-----LCVISVFMFGSPDLDSVSPFEDIRSRFIOGESITP 420
Db 374 SGMFVCAIFMSVGLVLLNFKFMSYVSMIAIF-----LFVSFFE-----IGP 415
Qy 421 TKIPEITTEIYMSNGSANSANVPETSPESVLIIS-----PRPALAIAAFSNWTCNFVALCFQVIADFCGPYVFFL 458
Db 416 GPIFWFVWAEFBSQ-----PRPALAIAAFSNWTCNFVALCFQVIADFCGPYVFFL 468
Qy 459 FAGVIAA 465
Db 469 FAGVLLA 475

RESULT 9
AI3443
Na+/H+ antiporter 1 / proteinase IV (EC 3.4.-.-) [imported] - Brucella melitensis (strain
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AI3443
R:DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujter, C.; Los, T.; Ivanova,
; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letese
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:1175668
A:Accession: AI3443
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-763 <KUR>
A:Cross-references: GB:AF008917; PIDN:AAU52716.1; PID:g17983545; GSPDB:GN00190
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI1535
A:Map position: I
C:Keywords: hydrolase

Query Match 4.1%; Score 120.5; DB 2; Length 763;
Best Local Similarity 21.5%; Pred. No. 0.26;
Matches 93; Conservative 55; Mismatches 150; Indels 135; Gaps 21;

Qy 55 GNSLLLTAVYGLVAGSV-----LVLGAI-IGDWVDKN-----ARLKVAQTS 95
Db 52 GITLMAAALALIVANSPPAQTYFDALHYIGPLSLAHWINDAUMAIFLLVGLLEIKREM 111
Qy 96 LVQNVQS-----VILCGIILMMVFLHKLHLLTMYHGWLTSCYIILITIAN 141
Db 112 LDGQLASWPNRMLPGIAAGGVILPAIFAVL---NHDNPAKLGWAVPSATDIAFALGV 168
Qy 142 IANLASTATAITQDRIWVVGEDRSKLANMNATIRRIQDLTNILAPMAVGQIMTFGSP 201
Db 169 LSLGSRAP-----SSLVFLATLAILDDLAAVVI-IAIFYTAETISMP 210
Qy 202 VIGCGFTSG-----WNLVSMCVVEVLLWKVYQKTPALAVKAGLKEEETELKQLN 250
Db 211 YLGAATIAAVLFVNRMDVVKLLPYLISAVILMFFVFNKSGVHATVAGVAA----- 262
Qy 251 LHKDTEPKPLEGTHMGVKD---SNIHELEHEOEPTCASQMAEPRTFRDGMVSYNQPV 307
Db 263 LMIFLKAP-----GRPDDMTSPLKLEH-----ALAKP-----VAFIVVPI 299
Qy 308 F-LAGMGLAP--LYMTVLGFCITTYAYTQGLSGSILSLMGASAITGIMGTVAFTWLR 364
Db 300 FGFANAGISFKGLEASVLG-DTLTLGI-----LLGLFLGKQ-----FGVFGAAWLA 344

Qy 365 RKCLGVRTGLISGLAQL-SCLILCVISVFMFGSPDLDSVSPFEDIRSRFIOGESITPTKI 423
Db 345 IKTLGAELKPGASVQVLYGVALIGGFTM-----SIFI-GLLSFPFSDL 387
Qy 424 PEITTEIYMSNGS 436
Db 388 MQTETKIGVLGSGS 400

RESULT 10
SI4144
hexose transport protein - Chlorella kessleri
C:Species: Chlorella kessleri
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 24-Sep-1999
C:Accession: SI4144; S07096
R:Wolf, K.; Tanner, W.; Sauer, N.
Curr. Genet. 19, 215-219, 1991
A:Title: The Chlorella H(+)/hexose cotransporter gene.
A:Reference number: SI4144; MUID:91330335; PMID:1868571
A:Accession: SI4144
A:Molecule type: DNA
A:Residues: 1-534 <WOL>
A:Cross-references: EMBL:X55349; NID:g18040; PIDN:CAA39036.1; PID:g18041
R:Sauer, N.; Tanner, W.
FEBS Lett. 259, 43-46, 1989
A:Title: The hexose carrier from Chlorella. cDNA cloning of a eucaryotic H(+)-cotranspor-
A:Reference number: S07096; MUID:90025536; PMID:2599110
A:Accession: S07096
A:Molecule type: mRNA
A:Residues: 1-58,60-208,'L',210-247,'R',249-534 <SAU>
A:Cross-references: EMBL:Y07520; NID:g18038; PIDN:CAA68813.1; PID:g18039
C:Genetics:
A:Gene: HUD1
A:Introns: 29/3; 60/3; 79/2; 114/3; 152/1; 194/2; 234/2; 274/3; 309/1; 366/3; 405/3; 465
C:Superfamily: Glucose transport protein
C:Keywords: sugar transport; transmembrane protein
F:26-46/Domain: transmembrane #status predicted <TM1>
F:88-108/Domain: transmembrane #status predicted <TM2>
F:117-137/Domain: transmembrane #status predicted <TM3>
F:145-165/Domain: transmembrane #status predicted <TM4>
F:172-192/Domain: transmembrane #status predicted <TM5>
F:202-222/Domain: transmembrane #status predicted <TM6>
F:288-308/Domain: transmembrane #status predicted <TM7>
F:323-344/Domain: transmembrane #status predicted <TM8>
F:353-373/Domain: transmembrane #status predicted <TM9>
F:388-408/Domain: transmembrane #status predicted <TM10>
F:427-447/Domain: transmembrane #status predicted <TM11>
F:457-477/Domain: transmembrane #status predicted <TM12>

Query Match 4.0%; Score 118.5; DB 2; Length 534;
Best Local Similarity 19.4%; Pred. No. 0.24;
Matches 123; Conservative 90; Mismatches 213; Indels 207; Gaps 29;

Qy 5 GDHNRQRCGCGSLADYLTGAKFLYLHLSLSTWGRMWHFAVSVFLVE----- 52
Db 17 GDRV-----GGITVTVVWVAFWAAACGGLLLGVDNGVTGCVSLEAFKKFFPDWAKQ 70
Qy 53 -----LYGNSLLLTAVYGLVAGSVLVLGAIIGDWVDKNARLKVAQ----- 93
Db 71 EVHEDSPYCTYDNAKLQLFVSSLFLAG---LVSCLFASWITRNWGRKVTWIGIGAPFVAG 127
Qy 94 --TSLVVQNVSVILCGIILM-----MVFLHKLHLLTMYHGWLTSCYIILITIA-N 141
Db 128 GLVNFAQDMAMLIIVGRVLLGVFGVGLGSQVVPQYLSSEVAFPSHRGMLNIGYQLFVTIGIL 187
Qy 142 IANLASTATAITQDRIWVVGEDRSKLANMNATIRRIQDLTNILAPMAV-GQIMTFGS 200
Db 188 IAGLVNAV-----RW-----ENGRLLS-----LGPAAAPGAILFLGS 221
Qy 201 PVTCGFIISGNLVSMCVVEVLLWKVYQKTPALAVKAGLKEEETELKQLNLRKDTPEKPL 260
Db 222 LVL-----PESPNFLVEKGTKEGREVLQ-----KL 247

Db 459 MSKFGFVSLSLERRAARFYIFNLINVLGSGVITGSAFEQDLSFL-----KOSAKEIP 513
Qy 444 ETPSPVPIISVSL-----FAGVIA--ARIGLWSFDLTWTQLLQNVIESERGIING 494
Db 514 KTVGVAIPIKATFFITYIMDGWAGIAGEILRLKPLIFFHKNLSLVKTEKDRBAMMPG 573
Qy 495 VONSNNY-----LDDLHFIMVILAPNPEAFGLLVLSVSVFAMGHIMY 538
Db 574 ---QINHYATBPRIQLYELGLGVAP---VTPVLLPFIIFALAYLVP 616

RESULT 13
C90446
permease, multidrug resistance protein [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 06-Jan-2003
C:Accession: C90446
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, R.
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A39139
A:Accession: C90446
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-469 <KUR>
A:Cross-references: GB:AB006641; NID:gl3816035; PIDN:AAK42826.1; GSPDB:GN00155
C:Genetics:
A:Gene: SSO2716
C:Superfamily: multidrug-efflux transporter

Query Match 4.0%; Score 116; DB 2; Length 469;
Best Local Similarity 22.7%; Pred. No. 0.32;
Matches 88; Conservative 64; Mismatches 143; Indels 92; Gaps 21;
Qy 68 VAGSVLVL--GAIIGDWDVKNARLVAQTSLVQVNSVILGCIIMVMVFLHKHLLTYH 125
Db 105 IGGSMVLNANSSIVADVPPNRRKAYGITSIGNIGA-LVGIVLGV-----LITFF 156
Qy 126 GWLTSCHILITITIANLANLASTATAITIQDRIQVNVAGEDRSKLANMNAIRIDQTN 185
Db 157 GWQ-----YIFINV-----PIGI-----VAVILGV--MNIKDKINKVTKLMLGA 195
Qy 186 ILAPMAVQIMTEGSPVIGCGFISGNLVSMC-----VEYVLLMKVYQKTPALAVKAGLK 240
Db 196 ILLGLSLA-LISLSIMFTAASGISTDNIELALGVALIPFLNRETRSKYPILNKI--- 251
Qy 241 EETELKQNLHKDTEPKPLEG-----THLMGVKDSNIHELEHEQEPTCASOMA 289
Db 252 -----FKIRLLTYSILANFLOGIGLSLSFLIMYLOQVR--GLSPDLSLLITPGYVIA 304
Qy 290 EPPFTFRDGVSYNQPVFLAGMGLAFLYMT-VLGFDCITTTGYAYTOGLSGSILSMGA 348
Db 305 SILAPFM-GRVADRGKPGIVAGIGLMFTIFMILYFFLLTPTDY-----VLIVGI 354
Qy 349 SAITGIMGTAVFTWLRKCGLV-----RTGLISGLAQ-----LSCLILCVISVMPGSP 398
Db 355 SAITGI-GSAMP-WPSNSTAMFHAPKEYGVSGLSRLTGNIGTILSYLSIVVATLSI 412
Qy 399 DLSVSPPE-----DIRSRITQ 415
Db 413 PRSVA-FEIFTGTTLNGDVSFTVNG 438

RESULT 14
E83686
hypothetical protein BH0293 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: E83686
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai

Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; NUID:20512582; PMID:11058132
A:Accession: E83686
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-418 <STO>
A:Cross-references: GB:AP001508; GB:BA000004; NID:gi0172890; PIDN:BA04012.1; GSPDB:GN00155
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH0293

Query Match 3.9%; Score 115.5; DB 2; Length 418;
Best Local Similarity 22.7%; Pred. No. 0.3;
Matches 112; Conservative 71; Mismatches 177; Indels 133; Gaps 27;
Qy 17 LADYLTSAKFL-----YLGHSLSWTGDRMWFHFAVSFVLVELYGNLSLLTAVYGLVVAG 70
Db 17 LSSLRNQKFLLLWVASIFSGLSFS-----IYFLTKSYVYVDELGLKTLFGLVLMAMTVP 71
Qy 71 SVLVGALIGDWDVKNARLVAQTSLVQVNSVILGCIIMVMVFLHKHLLTYHGWLT 130
Db 72 RVLLM-MIGGVLDPRFRRTIMPCSLITR--SVLLFGMFL-----QFLGVINIWTLI 121
Qy 131 SCYILIITI-----ANIANLASTATAITIQDRIQVNVVAGEDRSKLANMNAIRIDQTN 184
Db 122 GFAPCIGILDAPFPANVAILPS-----IVKKD-----HLTRANSFLOSTNHLT 165
Qy 185 NILAPMAVQIMTEGSPVIGCGFISGNLVSMC-----VEYVLLMKVYQKTPALAVKAGLK 240
Db 166 LIIGPMIAGMLIITTA-----YEIFVTIMCLLIIGAILIRV--KEPF----- 207
Qy 241 EETELKQNLHKDTEPKPLEGTHLMGVKQ--SNIHELEHEQEPTCASOMAEPTFRDG 298
Db 208 -DKPETKRL-----PSPL-----KLTGYSYLRH--SPVLLSLM----- 239
Qy 299 WVSYYNQPVFLAGMGLAFLYMTV---LGFDCTTTGYAYTOGLSGSILSMGASAITGIM 355
Db 240 WMSVFVN-FFITPGTWAIPLIVELRKGPDPLELSYLESTIAGLILGAIM-----IGII 293
Qy 356 GTVAPTWLRKCGL-----VRTGLISGLAQSLCLILCVISVMPGSPDLSPVSPFEDI 408
Db 294 N-----LRSKRGLIIIRLTVMGVALLEGQVSVLWQAIFVLAAILGWGISIG-----DI 342
Qy 409 RSRFIQGESITPKIPITTEIYMSNGSNSANIVPETSPEVPI-----TSVSLIFAG-- 461
Db 343 PARAVVOEK---TDFDKIGRWGMGMSTASSGLIPLSYGMTSVALSLSLSSILLYCGSF 399
Qy 462 -VIAARIGLWSFD 473
Db 400 VILFSPFVLWKFD 412

RESULT 15
T26524
hypothetical protein Y18D10A.23 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T26524
R;Harris, B.
submitted to the EMBL Data Library, December 1998
A:Reference number: Z20226
A:Accession: T26524
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-503 <WIL>
A:Cross-references: EMBL:AL034393; PIDN:CAA22315.1; CESP:Y18D10A.23
A:Experimental source: clone Y18D10A
C:Genetics:
A:Gene: CESP:Y18D10A.23
A:Introns: 32/3; 98/3; 271/1; 340/1; 423/2; 441/1
Query Match 3.9%; Score 115; DB 2; Length 503;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 18:24:46 ; Search time 19 Seconds
(without alignments)
1413.276 Million cell updates/sec

Title: US-09-715-927-6

Perfect score: 2929

Sequence: 1 MTRAGDHNRRQCGCGSLADY.....ACGPDKEVRKENQANTSW 571

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	126.5	4.3	523	1	GTR2_MOUSE
2	123.5	4.2	524	1	GTR2_HUMAN
3	118.5	4.0	534	1	HUPI_CHLKE
4	113.5	3.9	482	1	YF1G_BACSU
5	113.5	3.9	522	1	GTR2_RAT
6	112.5	3.8	435	1	YVBO_BACSU
7	109.5	3.7	461	1	CSEC_BACSU
8	108.5	3.7	400	1	RFAY_XANCP
9	108	3.7	618	1	SL55_RAT
10	106.5	3.6	492	1	YD19_METUA
11	106	3.6	547	1	YQJ1_ECOLI
12	104.5	3.6	502	1	GTR5_RAT
13	103.5	3.5	455	1	AZAC_CAVPO
14	103.5	3.5	533	1	GTR2_CHICK
15	103	3.5	358	1	HH2R_RAT
16	103	3.5	634	1	SL56_RAT
17	103	3.5	718	1	SL53_HUMAN
18	103	3.5	1043	1	P11D_MOUSE
19	102.5	3.5	394	1	YWFF_BACSU
20	102.5	3.5	429	1	ASBA_YEREN
21	102	3.5	547	1	MERA_STARP
22	101.5	3.5	458	1	AZAC_MOUSE
23	101.5	3.5	458	1	AZAC_RAT
24	101.5	3.5	632	1	AFUB_HAEIN
25	101.5	3.5	739	1	DDID_HUMAN
26	101.5	3.5	971	1	Y277_MYCPN
27	101	3.4	457	1	UIWB_ECOLI
28	100.5	3.4	411	1	CIW2_MOUSE
29	100.5	3.4	552	1	YH81_MYCTU
30	99.5	3.4	460	1	NU4M_SQUAC
31	99.5	3.4	496	1	GTR3_HUMAN
32	98.5	3.4	472	1	EXUT_ECOLI
33	98.5	3.4	822	1	VGLE_HSV7J

RESULT 1

GTR2_MOUSE

ID GTR2_MOUSE STANDARD; PRT; 523 AA.

AC P14246; Q9DBA7;

DT 01-JAN-1990 (Rel. 13, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Solute carrier family 2, facilitated glucose transporter, member 2

DE (Glucose transporter type 2, liver).

GN SLC2A2 OR GLUT2 OR GLUT-2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/c; TISSUE=Liver;

RX MEDLINE=90098776; PubMed=2602116;

RA Suzue K., Lodish H.F., Thorens B.;

RT "Sequence of the mouse liver glucose transporter.";

RL Nucleic Acids Res. 17:10099-10099(1989).

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6; TISSUE=Liver;

RX MEDLINE=89366666; PubMed=2771649;

RA Asano T., Shibasaki Y., Lin J.L., Akanuma Y., Takaku F., Oka Y.;

RT "The nucleotide sequence of cDNA for a mouse liver-type glucose transporter protein.";

RL Nucleic Acids Res. 17:6386-6386(1989).

[3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Liver;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Mofelli D., Bojunga N., Carninci P., de Bonaudo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

[4]

RP SEQUENCE OF 384-496 FROM N.A.

RX MEDLINE=92111400; PubMed=1765007;

P18825 homo sapien
P03819 escherichia
Q48624 leuconostoc
P53793 bos taurus
P97046 lactococcus
P25336 saccharomyc
Q91is4 rattus norv
Q12564 ampelomyces
P97292 mus musculus
O79881 sus scrofa
O87712 coxiella bu
Q9m3j4 spinacia ol

RA Hogan A., Heyner S., Charlton M.J., Copeland N.G., Gilbert D.J.,
 RA Jenkins N.A., Thorens B., Schultz G.A.; Gilbert D.J.,
 RL "Glucose transporter gene expression in early mouse embryos."
 RL Development 113:363-372(1991).
 [5]
 RN TISSUE SPECIFICITY.
 RC STRAIN=C57BL/6;
 RC MEDLINE=93170163; PubMed=1289053;
 RA Smith D.E., Gridley T.;
 RT "Differential screening of a PCR-generated mouse embryo cDNA library:
 RT glucose transporters are differentially expressed in early
 RT postimplantation mouse embryos."
 RL Development 116:555-561(1992).
 CC -1- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY
 CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
 CC MEMBRANE OF HEPATOCYTES AND IS RESPONSIBLE FOR UPTAKE OF GLUCOSE
 CC BY THE BETA CELLS; MAY COMPRISE PART OF THE GLUCOSE-SENSING
 CC MECHANISM OF THE BETA CELL. MAY ALSO PARTICIPATE WITH THE
 CC NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCELLULAR TRANSPORT OF
 CC GLUCOSE IN THE SMALL INTESTINE AND KIDNEY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: IN EMBRYO, EXPRESSED IN ENDODERM LAYER OF
 CC YOLK SAC AND LIVER PRIMORDIUM.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
 CC TRANSPORTER SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X16986; CAA34855.1; -
 CC EMBL; X15684; CAA33719.1; -
 CC EMBL; S77926; AAB20847.1; -
 CC EMBL; AK005068; BAB23792.1; -
 CC F01; S06920; S06920.
 CC MGD; MGI:1095438; Slc2a2.
 CC InterPro; IPR007114; MFS.
 CC InterPro; IPR005828; Sub transporter.
 CC InterPro; IPR005829; Sug_transp.
 CC InterPro; IPR003663; Sugar_transp.
 CC Pfam; PF00083; sugar_tr; 1.
 CC PRINTS; PR00171; SUGRTNSPOT.
 CC TIGRFAMs; TIGR00879; SP; 1.
 CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Transmembrane; Sugar transporter; Transport; Glycoprotein;
 KW Multigene family.
 FT DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 11 31 1 (POTENTIAL).
 FT DOMAIN 32 97 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 98 118 2 (POTENTIAL).
 FT TRANSMEM 119 126 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 127 147 3 (POTENTIAL).
 FT TRANSMEM 148 157 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 158 178 4 (POTENTIAL).
 FT TRANSMEM 179 186 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 187 207 5 (POTENTIAL).
 FT DOMAIN 208 216 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 217 237 6 (POTENTIAL).
 FT DOMAIN 238 258 7 (POTENTIAL).
 FT TRANSMEM 259 277 8 (POTENTIAL).
 FT DOMAIN 278 297 9 (POTENTIAL).
 FT TRANSMEM 298 317 10 (POTENTIAL).
 FT DOMAIN 318 337 11 (POTENTIAL).
 FT TRANSMEM 338 358 12 (POTENTIAL).
 FT DOMAIN 359 377 13 (POTENTIAL).
 FT TRANSMEM 378 397 14 (POTENTIAL).
 FT DOMAIN 398 417 15 (POTENTIAL).
 FT TRANSMEM 418 437 16 (POTENTIAL).
 FT DOMAIN 438 453 17 (POTENTIAL).
 FT TRANSMEM 454 473 18 (POTENTIAL).
 FT TRANSMEM 474 493 19 (POTENTIAL).
 FT TRANSMEM 494 513 20 (POTENTIAL).
 FT TRANSMEM 514 533 21 (POTENTIAL).
 FT TRANSMEM 534 553 22 (POTENTIAL).
 FT TRANSMEM 554 573 23 (POTENTIAL).
 FT TRANSMEM 574 593 24 (POTENTIAL).
 FT TRANSMEM 594 613 25 (POTENTIAL).
 FT TRANSMEM 614 633 26 (POTENTIAL).
 FT TRANSMEM 634 653 27 (POTENTIAL).
 FT TRANSMEM 654 673 28 (POTENTIAL).
 FT TRANSMEM 674 693 29 (POTENTIAL).
 FT TRANSMEM 694 713 30 (POTENTIAL).
 FT TRANSMEM 714 733 31 (POTENTIAL).
 FT TRANSMEM 734 753 32 (POTENTIAL).
 FT TRANSMEM 754 773 33 (POTENTIAL).
 FT TRANSMEM 774 793 34 (POTENTIAL).
 FT TRANSMEM 794 813 35 (POTENTIAL).
 FT TRANSMEM 814 833 36 (POTENTIAL).
 FT TRANSMEM 834 853 37 (POTENTIAL).
 FT TRANSMEM 854 873 38 (POTENTIAL).
 FT TRANSMEM 874 893 39 (POTENTIAL).
 FT TRANSMEM 894 913 40 (POTENTIAL).
 FT TRANSMEM 914 933 41 (POTENTIAL).
 FT TRANSMEM 934 953 42 (POTENTIAL).
 FT TRANSMEM 954 973 43 (POTENTIAL).
 FT TRANSMEM 974 993 44 (POTENTIAL).
 FT TRANSMEM 994 1013 45 (POTENTIAL).
 FT TRANSMEM 1014 1033 46 (POTENTIAL).
 FT TRANSMEM 1034 1053 47 (POTENTIAL).
 FT TRANSMEM 1054 1073 48 (POTENTIAL).
 FT TRANSMEM 1074 1093 49 (POTENTIAL).
 FT TRANSMEM 1094 1113 50 (POTENTIAL).
 FT TRANSMEM 1114 1133 51 (POTENTIAL).
 FT TRANSMEM 1134 1153 52 (POTENTIAL).
 FT TRANSMEM 1154 1173 53 (POTENTIAL).
 FT TRANSMEM 1174 1193 54 (POTENTIAL).
 FT TRANSMEM 1194 1213 55 (POTENTIAL).
 FT TRANSMEM 1214 1233 56 (POTENTIAL).
 FT TRANSMEM 1234 1253 57 (POTENTIAL).
 FT TRANSMEM 1254 1273 58 (POTENTIAL).
 FT TRANSMEM 1274 1293 59 (POTENTIAL).
 FT TRANSMEM 1294 1313 60 (POTENTIAL).
 FT TRANSMEM 1314 1333 61 (POTENTIAL).
 FT TRANSMEM 1334 1353 62 (POTENTIAL).
 FT TRANSMEM 1354 1373 63 (POTENTIAL).
 FT TRANSMEM 1374 1393 64 (POTENTIAL).
 FT TRANSMEM 1394 1413 65 (POTENTIAL).
 FT TRANSMEM 1414 1433 66 (POTENTIAL).
 FT TRANSMEM 1434 1453 67 (POTENTIAL).
 FT TRANSMEM 1454 1473 68 (POTENTIAL).
 FT TRANSMEM 1474 1493 69 (POTENTIAL).
 FT TRANSMEM 1494 1513 70 (POTENTIAL).
 FT TRANSMEM 1514 1533 71 (POTENTIAL).
 FT TRANSMEM 1534 1553 72 (POTENTIAL).
 FT TRANSMEM 1554 1573 73 (POTENTIAL).
 FT TRANSMEM 1574 1593 74 (POTENTIAL).
 FT TRANSMEM 1594 1613 75 (POTENTIAL).
 FT TRANSMEM 1614 1633 76 (POTENTIAL).
 FT TRANSMEM 1634 1653 77 (POTENTIAL).
 FT TRANSMEM 1654 1673 78 (POTENTIAL).
 FT TRANSMEM 1674 1693 79 (POTENTIAL).
 FT TRANSMEM 1694 1713 80 (POTENTIAL).
 FT TRANSMEM 1714 1733 81 (POTENTIAL).
 FT TRANSMEM 1734 1753 82 (POTENTIAL).
 FT TRANSMEM 1754 1773 83 (POTENTIAL).
 FT TRANSMEM 1774 1793 84 (POTENTIAL).
 FT TRANSMEM 1794 1813 85 (POTENTIAL).
 FT TRANSMEM 1814 1833 86 (POTENTIAL).
 FT TRANSMEM 1834 1853 87 (POTENTIAL).
 FT TRANSMEM 1854 1873 88 (POTENTIAL).
 FT TRANSMEM 1874 1893 89 (POTENTIAL).
 FT TRANSMEM 1894 1913 90 (POTENTIAL).
 FT TRANSMEM 1914 1933 91 (POTENTIAL).
 FT TRANSMEM 1934 1953 92 (POTENTIAL).
 FT TRANSMEM 1954 1973 93 (POTENTIAL).
 FT TRANSMEM 1974 1993 94 (POTENTIAL).
 FT TRANSMEM 1994 2013 95 (POTENTIAL).
 FT TRANSMEM 2014 2033 96 (POTENTIAL).
 FT TRANSMEM 2034 2053 97 (POTENTIAL).
 FT TRANSMEM 2054 2073 98 (POTENTIAL).
 FT TRANSMEM 2074 2093 99 (POTENTIAL).
 FT TRANSMEM 2094 2113 100 (POTENTIAL).
 FT TRANSMEM 2114 2133 101 (POTENTIAL).
 FT TRANSMEM 2134 2153 102 (POTENTIAL).
 FT TRANSMEM 2154 2173 103 (POTENTIAL).
 FT TRANSMEM 2174 2193 104 (POTENTIAL).
 FT TRANSMEM 2194 2213 105 (POTENTIAL).
 FT TRANSMEM 2214 2233 106 (POTENTIAL).
 FT TRANSMEM 2234 2253 107 (POTENTIAL).
 FT TRANSMEM 2254 2273 108 (POTENTIAL).
 FT TRANSMEM 2274 2293 109 (POTENTIAL).
 FT TRANSMEM 2294 2313 110 (POTENTIAL).
 FT TRANSMEM 2314 2333 111 (POTENTIAL).
 FT TRANSMEM 2334 2353 112 (POTENTIAL).
 FT TRANSMEM 2354 2373 113 (POTENTIAL).
 FT TRANSMEM 2374 2393 114 (POTENTIAL).
 FT TRANSMEM 2394 2413 115 (POTENTIAL).
 FT TRANSMEM 2414 2433 116 (POTENTIAL).
 FT TRANSMEM 2434 2453 117 (POTENTIAL).
 FT TRANSMEM 2454 2473 118 (POTENTIAL).
 FT TRANSMEM 2474 2493 119 (POTENTIAL).
 FT TRANSMEM 2494 2513 120 (POTENTIAL).
 FT TRANSMEM 2514 2533 121 (POTENTIAL).
 FT TRANSMEM 2534 2553 122 (POTENTIAL).
 FT TRANSMEM 2554 2573 123 (POTENTIAL).
 FT TRANSMEM 2574 2593 124 (POTENTIAL).
 FT TRANSMEM 2594 2613 125 (POTENTIAL).
 FT TRANSMEM 2614 2633 126 (POTENTIAL).
 FT TRANSMEM 2634 2653 127 (POTENTIAL).
 FT TRANSMEM 2654 2673 128 (POTENTIAL).
 FT TRANSMEM 2674 2693 129 (POTENTIAL).
 FT TRANSMEM 2694 2713 130 (POTENTIAL).
 FT TRANSMEM 2714 2733 131 (POTENTIAL).
 FT TRANSMEM 2734 2753 132 (POTENTIAL).
 FT TRANSMEM 2754 2773 133 (POTENTIAL).
 FT TRANSMEM 2774 2793 134 (POTENTIAL).
 FT TRANSMEM 2794 2813 135 (POTENTIAL).
 FT TRANSMEM 2814 2833 136 (POTENTIAL).
 FT TRANSMEM 2834 2853 137 (POTENTIAL).
 FT TRANSMEM 2854 2873 138 (POTENTIAL).
 FT TRANSMEM 2874 2893 139 (POTENTIAL).
 FT TRANSMEM 2894 2913 140 (POTENTIAL).
 FT TRANSMEM 2914 2933 141 (POTENTIAL).
 FT TRANSMEM 2934 2953 142 (POTENTIAL).
 FT TRANSMEM 2954 2973 143 (POTENTIAL).
 FT TRANSMEM 2974 2993 144 (POTENTIAL).
 FT TRANSMEM 2994 3013 145 (POTENTIAL).
 FT TRANSMEM 3014 3033 146 (POTENTIAL).
 FT TRANSMEM 3034 3053 147 (POTENTIAL).
 FT TRANSMEM 3054 3073 148 (POTENTIAL).
 FT TRANSMEM 3074 3093 149 (POTENTIAL).
 FT TRANSMEM 3094 3113 150 (POTENTIAL).
 FT TRANSMEM 3114 3133 151 (POTENTIAL).
 FT TRANSMEM 3134 3153 152 (POTENTIAL).
 FT TRANSMEM 3154 3173 153 (POTENTIAL).
 FT TRANSMEM 3174 3193 154 (POTENTIAL).
 FT TRANSMEM 3194 3213 155 (POTENTIAL).
 FT TRANSMEM 3214 3233 156 (POTENTIAL).
 FT TRANSMEM 3234 3253 157 (POTENTIAL).
 FT TRANSMEM 3254 3273 158 (POTENTIAL).
 FT TRANSMEM 3274 3293 159 (POTENTIAL).
 FT TRANSMEM 3294 3313 160 (POTENTIAL).
 FT TRANSMEM 3314 3333 161 (POTENTIAL).
 FT TRANSMEM 3334 3353 162 (POTENTIAL).
 FT TRANSMEM 3354 3373 163 (POTENTIAL).
 FT TRANSMEM 3374 3393 164 (POTENTIAL).
 FT TRANSMEM 3394 3413 165 (POTENTIAL).
 FT TRANSMEM 3414 3433 166 (POTENTIAL).
 FT TRANSMEM 3434 3453 167 (POTENTIAL).
 FT TRANSMEM 3454 3473 168 (POTENTIAL).
 FT TRANSMEM 3474 3493 169 (POTENTIAL).
 FT TRANSMEM 3494 3513 170 (POTENTIAL).
 FT TRANSMEM 3514 3533 171 (POTENTIAL).
 FT TRANSMEM 3534 3553 172 (POTENTIAL).
 FT TRANSMEM 3554 3573 173 (POTENTIAL).
 FT TRANSMEM 3574 3593 174 (POTENTIAL).
 FT TRANSMEM 3594 3613 175 (POTENTIAL).
 FT TRANSMEM 3614 3633 176 (POTENTIAL).
 FT TRANSMEM 3634 3653 177 (POTENTIAL).
 FT TRANSMEM 3654 3673 178 (POTENTIAL).
 FT TRANSMEM 3674 3693 179 (POTENTIAL).
 FT TRANSMEM 3694 3713 180 (POTENTIAL).
 FT TRANSMEM 3714 3733 181 (POTENTIAL).
 FT TRANSMEM 3734 3753 182 (POTENTIAL).
 FT TRANSMEM 3754 3773 183 (POTENTIAL).
 FT TRANSMEM 3774 3793 184 (POTENTIAL).
 FT TRANSMEM 3794 3813 185 (POTENTIAL).
 FT TRANSMEM 3814 3833 186 (POTENTIAL).
 FT TRANSMEM 3834 3853 187 (POTENTIAL).
 FT TRANSMEM 3854 3873 188 (POTENTIAL).
 FT TRANSMEM 3874 3893 189 (POTENTIAL).
 FT TRANSMEM 3894 3913 190 (POTENTIAL).
 FT TRANSMEM 3914 3933 191 (POTENTIAL).
 FT TRANSMEM 3934 3953 192 (POTENTIAL).
 FT TRANSMEM 3954 3973 193 (POTENTIAL).
 FT TRANSMEM 3974 3993 194 (POTENTIAL).
 FT TRANSMEM 3994 4013 195 (POTENTIAL).
 FT TRANSMEM 4014 4033 196 (POTENTIAL).
 FT TRANSMEM 4034 4053 197 (POTENTIAL).
 FT TRANSMEM 4054 4073 198 (POTENTIAL).
 FT TRANSMEM 4074 4093 199 (POTENTIAL).
 FT TRANSMEM 4094 4113 200 (POTENTIAL).
 FT TRANSMEM 4114 4133 201 (POTENTIAL).
 FT TRANSMEM 4134 4153 202 (POTENTIAL).
 FT TRANSMEM 4154 4173 203 (POTENTIAL).
 FT TRANSMEM 4174 4193 204 (POTENTIAL).
 FT TRANSMEM 4194 4213 205 (POTENTIAL).
 FT TRANSMEM 4214 4233 206 (POTENTIAL).
 FT TRANSMEM 4234 4253 207 (POTENTIAL).
 FT TRANSMEM 4254 4273 208 (POTENTIAL).
 FT TRANSMEM 4274 4293 209 (POTENTIAL).
 FT TRANSMEM 4294 4313 210 (POTENTIAL).
 FT TRANSMEM 4314 4333 211 (POTENTIAL).
 FT TRANSMEM 4334 4353 212 (POTENTIAL).
 FT TRANSMEM 4354 4373 213 (POTENTIAL).
 FT TRANSMEM 4374 4393 214 (POTENTIAL).
 FT TRANSMEM 4394 4413 215 (POTENTIAL).
 FT TRANSMEM 4414 4433 216 (POTENTIAL).
 FT TRANSMEM 4434 4453 217 (POTENTIAL).
 FT TRANSMEM 4454 4473 218 (POTENTIAL).
 FT TRANSMEM 4474 4493 219 (POTENTIAL).
 FT TRANSMEM 4494 4513 220 (POTENTIAL).
 FT TRANSMEM 4514 4533 221 (POTENTIAL).
 FT TRANSMEM 4534 4553 222 (POTENTIAL).
 FT TRANSMEM 4554 4573 223 (POTENTIAL).
 FT TRANSMEM 4574 4593 224 (POTENTIAL).
 FT TRANSMEM 4594 4613 225 (POTENTIAL).
 FT TRANSMEM 4614 4633 226 (POTENTIAL).
 FT TRANSMEM 4634 4653 227 (POTENTIAL).
 FT TRANSMEM 4654 4673 228 (POTENTIAL).
 FT TRANSMEM 4674 4693 229 (POTENTIAL).
 FT TRANSMEM 4694 4713 230 (POTENTIAL).
 FT TRANSMEM 4714 4733 231 (POTENTIAL).
 FT TRANSMEM 4734 4753 232 (POTENTIAL).
 FT TRANSMEM 4754 4773 233 (POTENTIAL).
 FT TRANSMEM 4774 4793 234 (POTENTIAL).
 FT TRANSMEM 4794 4813 235 (POTENTIAL).
 FT TRANSMEM 4814 4833 236 (POTENTIAL).
 FT TRANSMEM 4834 4853 237 (POTENTIAL).
 FT TRANSMEM 4854 4873 238 (POTENTIAL).
 FT TRANSMEM 4874 4893 239 (POTENTIAL).
 FT TRANSMEM 4894 4913 240 (POTENTIAL).
 FT TRANSMEM 4914 4933 241 (POTENTIAL).
 FT TRANSMEM 4934 4953 242 (POTENTIAL).
 FT TRANSMEM 4954 4973 243 (POTENTIAL).
 FT TRANSMEM 4974 4993 244 (POTENTIAL).
 FT TRANSMEM 4994 5013 245 (POTENTIAL).
 FT TRANSMEM 5014 5033 246 (POTENTIAL).
 FT TRANSMEM 5034 5053 247 (POTENTIAL).
 FT TRANSMEM 5054 5073 248 (POTENTIAL).
 FT TRANSMEM 5074 5093 249 (POTENTIAL).
 FT TRANSMEM 5094 5113 250 (POTENTIAL).
 FT TRANSMEM 5114 5133 251 (POTENTIAL).
 FT TRANSMEM 5134 5153 252 (POTENTIAL).
 FT TRANSMEM 5154 5173 253 (POTENTIAL).
 FT TRANSMEM 5174 5193 254 (POTENTIAL).
 FT TRANSMEM 5194 5213 255 (POTENTIAL).
 FT TRANSMEM 5214 5233 256 (POTENTIAL).
 FT TRANSMEM 5234 5253 257 (POTENTIAL).
 FT TRANSMEM 5254 5273 258 (POTENTIAL).
 FT TRANSMEM 5274 5293 259 (POTENTIAL).
 FT TRANSMEM 5294 5313 260 (POTENTIAL).
 FT TRANSMEM 5314 5333 261 (POTENTIAL).
 FT TRANSMEM 5334 5353 262 (POTENTIAL).
 FT TRANSMEM 5354 5373 263 (POTENTIAL).
 FT TRANSMEM 5374 5393 264 (POTENTIAL).
 FT TRANSMEM 5394 5413 265 (POTENTIAL).
 FT TRANSMEM 5414 5433 266 (POTENTIAL).
 FT TRANSMEM 5434 5453 267 (POTENTIAL).
 FT TRANSMEM 5454 5473 268 (POTENTIAL).
 FT TRANSMEM 5474 5493 269 (POTENTIAL).
 FT TRANSMEM 5494 5513 270 (POTENTIAL).
 FT TRANSMEM 5514 5533 271 (POTENTIAL).
 FT TRANSMEM 5534 5553 272 (POTENTIAL).
 FT TRANSMEM 5554 5573 273 (POTENTIAL).
 FT TRANSMEM 5574 5593 274 (POTENTIAL).
 FT TRANSMEM 5594 5613 275 (POTENTIAL).
 FT TRANSMEM 5614 5633 276 (POTENTIAL).
 FT TRANSMEM 5634 5653 277 (POTENTIAL).
 FT TRANSMEM 5654 5673 278 (POTENTIAL).
 FT TRANSMEM 5674 5693 279 (POTENTIAL).
 FT TRANSMEM 5694 5713 280 (POTENTIAL).
 FT TRANSMEM 5714 5733 281 (POTENTIAL).
 FT TRANSMEM 5734 5753 282 (POTENTIAL).
 FT TRANSMEM 5754 5773 283 (POTENTIAL).
 FT TRANSMEM 5774 5793 284 (POTENTIAL).
 FT TRANSMEM 5794 5813 285 (POTENTIAL).
 FT TRANSMEM 5814 5833 286 (POTENTIAL).
 FT TRANSMEM 5834 5853 287 (POTENTIAL).
 FT TRANSMEM 5854 5873 288 (POTENTIAL).
 FT TRANSMEM 5874 5893 289 (POTENTIAL).
 FT TRANSMEM 5894 5913 290 (POTENTIAL).
 FT TRANSMEM 5914 5933 291 (POTENTIAL).
 FT TRANSMEM 5934 5953 292 (POTENTIAL).
 FT TRANSMEM 5954 5973 293 (POTENTIAL).
 FT TRANSMEM 5974 5993 294 (POTENTIAL).
 FT TRANSMEM 5994 6013 295 (POTENTIAL).
 FT TRANSMEM 6014 6033 296 (POTENTIAL).
 FT TRANSMEM 6034 6053 297 (POTENTIAL).
 FT TRANSMEM 6054 6073 298 (POTENTIAL).
 FT TRANSMEM 6074 6093 299 (POTENTIAL).
 FT TRANSMEM 6094 6113 300 (POTENTIAL).
 FT TRANSMEM 6114 6133 301 (POTENTIAL).
 FT TRANSMEM 6134 6153 302 (POTENTIAL).
 FT TRANSMEM 6154 6173 303 (POTENTIAL).
 FT TRANSMEM 6174 6193 304 (POTENTIAL).
 FT TRANSMEM 6194 6213 305 (POTENTIAL).
 FT TRANSMEM 6214 6233 306 (POTENTIAL).
 FT TRANSMEM 6234 6253 307 (POTENTIAL).
 FT TRANSMEM 6254 6273 308 (POTENTIAL).
 FT TRANSMEM 6274 6293 309 (POTENTIAL).
 FT TRANSMEM 6294 6313 310 (POTENTIAL).
 FT TRANSMEM 6314 6333 311 (POTENTIAL).
 FT TRANSMEM 6334 6353 312 (POTENTIAL).
 FT TRANSMEM 6354 6373 313 (POTENTIAL).
 FT TRANSMEM 6374 6393 314 (POTENTIAL).
 FT TRANSMEM 6394 6413 315 (POTENTIAL).
 FT TRANSMEM 6414 6433 316 (POTENTIAL).
 FT TRANSMEM 6434 6453 317 (POTENTIAL).
 FT TRANSMEM 6454 6473 318 (POTENTIAL).
 FT TRANSMEM 6474 6493 319 (POTENTIAL).
 FT TRANSMEM 6494 6513 320 (POTENTIAL).
 FT TRANSMEM 6514 6533 321 (POTENTIAL).
 FT TRANSMEM 6534 6553 322 (POTENTIAL).
 FT TRANSMEM 6554 6573 323 (POTENTIAL).
 FT TRANSMEM 6574 6593 324 (POTENTIAL).
 FT TRANSMEM 6594 6613 325 (POTENTIAL).
 FT TRANSMEM 6614 6633 326 (POTENTIAL).
 FT TRANSMEM 6634 6653 327 (POTENTIAL).
 FT TRANSMEM 6654 6673 328 (POTENTIAL).
 FT TRANSMEM 6674 6693 329 (POTENTIAL).
 FT TRANSMEM 6694 6713 330 (POTENTIAL).
 FT TRANSMEM 6714 6733 331 (POTENTIAL).
 FT TRANSMEM 6734 6753 332 (POTENTIAL).
 FT TRANSMEM 6754 6773 333 (POTENTIAL).
 FT TRANSMEM 6774 6793 334 (POTENTIAL).
 FT TRANSMEM 6794 6813 335 (POTENTIAL).
 FT TRANSMEM 6814 6833 336 (POTENTIAL).
 FT TRANSMEM 6834 6853 337 (POTENTIAL).
 FT TRANSMEM 6854 6873 338 (POTENTIAL).
 FT TRANSMEM 6874 6893 339 (POTENTIAL).
 FT TRANSMEM 6894 6913 340 (POTENTIAL).
 FT TRANSMEM 6914 6933 341 (POTENTIAL).
 FT TRANSMEM 6934 6953 342 (POTENTIAL).
 FT TRANSMEM 6954 6973 343 (POTENTIAL).
 FT TRANSMEM 6974 6993 344 (POTENTIAL).
 FT TRANSMEM 6994 7013 345 (POTENTIAL).
 FT TRANSMEM 7014 7033 346 (POTENTIAL).
 FT TRANSMEM 7034 7053 347 (POTENTIAL).
 FT TRANSMEM 7054 7073 348 (POTENTIAL).
 FT TRANSMEM 7074 7093 349 (POTENTIAL).
 FT TRANSMEM 7094 7113 350 (POTENTIAL).
 FT TRANSMEM 7114 7133 351 (POTENTIAL).
 FT TRANSMEM 7134 7153 352 (POTENTIAL).
 FT TRANSMEM 7154 7173 353 (POTENTIAL).
 FT TRANSMEM 7174 7193 354 (POTENTIAL).
 FT TRANSMEM 7194 7213 355 (POTENTIAL).
 FT TRANSMEM 7214 7233 356 (POTENTIAL).
 FT TRANSMEM 7234 7253 357 (POTENTIAL).
 FT TRANSMEM 7254 7273 358 (POTENTIAL).
 FT TRANSMEM 7274 7293 359 (POTENTIAL).
 FT TRANSMEM 7294 7313 360 (POTENTIAL).
 FT TRANSMEM 7314 7333 361 (POTENTIAL).
 FT TRANSMEM 7334 7353 362 (POTENTIAL).
 FT TRANSMEM 7354 7373 363 (POTENTIAL).
 FT TRANSMEM 7374 7393 364 (POTENTIAL).
 FT TRANSMEM 7394 7413 365 (POTENTIAL).
 FT TRANSMEM 7414 7433 366 (POTENTIAL).
 FT TRANSMEM 7434 7453 367 (POTENTIAL).
 FT TRANSMEM 7454 7473 368 (POTENTIAL).
 FT TRANSMEM 7474 7493 369 (POTENTIAL).
 FT TRANSMEM 7494 7513 370 (POTENTIAL).
 FT TRANSMEM 7514 7533 371 (POTENTIAL).
 FT TRANSMEM 7534 7553 372 (POTENTIAL).
 FT TRANSMEM 7554 7573 373 (POTENTIAL).
 FT TRANSMEM 7574 7593 374 (POTENTIAL).
 FT TRANSMEM 7594 7613 375 (POTENTIAL).
 FT TRANSMEM 7614 7633 376 (POTENTIAL).
 FT TRANSMEM 7634 7653 377 (POTENTIAL).
 FT TRANSMEM 7654 7673 378 (POTENTIAL).
 FT TRANSMEM 7674 7693 379 (POTENTIAL).
 FT TRANSMEM 7694 7713 380 (POTENTIAL).
 FT TRANSMEM 7714 7733 381 (POTENTIAL).
 FT TRANSMEM 7734 7753 382 (POTENTIAL).
 FT TRANSMEM 7754 7773 383 (POTENTIAL).
 FT TRANSMEM 7774 7793 384 (POTENTIAL).
 FT TRANSMEM 7794 7813 385 (POTENTIAL).
 FT TRANSMEM 7814 7833 386 (POTENTIAL).
 FT TRANSMEM 7834 7853 387 (POTENTIAL).
 FT TRANSMEM 7854 7873 388 (POTENTIAL).
 FT TRANSMEM 7874 7893 389 (POTENTIAL).
 FT TRANSMEM 7894 7913 390 (POTENTIAL).
 FT TRANSMEM 7914 7933 391 (POTENTIAL).
 FT TRANSMEM 7934 7953 392 (POTENTIAL).
 FT TRANSMEM 7954 7973 393 (POTENTIAL).
 FT TRANSMEM 7974 7993 394 (POTENTIAL).
 FT TRANSMEM 7994 8013 395 (POTENTIAL).
 FT TRANSMEM 8014 8033 396 (POTENTIAL).
 FT TRANSMEM 8034 8053 397 (POTENTIAL).
 FT TRANSMEM 8054 8073 398 (POTENTIAL).
 FT TRANSMEM 8074 8093 399 (POTENTIAL).
 FT TRANSMEM 8094 8113 400 (POTENTIAL).
 FT TRANSMEM 8114 8133 401 (POTENTIAL).
 FT TRANSMEM 8134 8153 402 (POTENTIAL).
 FT TRANSMEM 8154 8173 403 (POTENTIAL).
 FT TRANSMEM 8174 8193 404 (POTENTIAL).
 FT TRANSMEM 8194 8213 405 (POTENTIAL).
 FT TRANSMEM 8214 8233 406 (POTENTIAL).
 FT TRANSMEM 8234 8253 407 (POTENTIAL).
 FT TRANSMEM 8254 8273 408 (POTENTIAL).
 FT TRANSMEM 8274 8293 409 (POTENTIAL).
 FT TRANSMEM 8294 8313 410 (POTENTIAL).
 FT TRANSMEM 8314 8333 411 (POTENTIAL).
 FT TRANSMEM 8334 8353 412 (POTENTIAL).
 FT TRANSMEM 8354 8373 413 (POTENTIAL).
 FT TRANSMEM 8374 8393 414 (POTENTIAL).
 FT TRANSMEM 839

CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; Y07520; CAA68813.1; -

DR EMBL; X55349; CAA39036.1; -

DR PIR; S14144; S14144

DR InterPro; IPR007114; MFS.

DR InterPro; IPR005828; Sug_transporter.

DR InterPro; IPR005829; Sub_transporter.

DR InterPro; IPR003663; Sugar transp.

DR Pfam; PF00083; sugar tr; 1.

DR PRINTS; PR00171; SUGRTRANSPORT.

DR TIGR; TIGR00879; SP; 1.

DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.

DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.

DR Transmembrane; Transp; Sugar transport; Symport; Repeat;

DR Multigene family.

FT DOMAIN 1 21 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 22 42 1 (POTENTIAL).

FT DOMAIN 43 87 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 88 108 2 (POTENTIAL).

FT DOMAIN 109 114 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 115 135 3 (POTENTIAL).

FT DOMAIN 136 144 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 145 165 4 (POTENTIAL).

FT DOMAIN 166 173 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 174 194 5 (POTENTIAL).

FT DOMAIN 195 204 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 205 225 6 (POTENTIAL).

FT DOMAIN 226 299 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 300 322 7 (POTENTIAL).

FT DOMAIN 323 328 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 329 349 8 (POTENTIAL).

FT DOMAIN 350 352 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 353 373 9 (POTENTIAL).

FT DOMAIN 374 387 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 388 408 10 (POTENTIAL).

FT DOMAIN 409 433 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 434 454 11 (POTENTIAL).

FT DOMAIN 455 456 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 457 477 12 (POTENTIAL).

FT DOMAIN 478 534 CYTOPLASMIC (POTENTIAL).

FT CONFLICT 60 MISSING (IN REF. 1).

FT CONFLICT 209 P -> L (IN REF. 1).

FT CONFLICT 248 C -> R (IN REF. 1).

SQ SEQUENCE 534 AA; 57522 MW; 3B23BE8710AA2237 CRC64;

Query Match 4.08; Score 118.5; DB 1; Length 534;

Best Local Similarity 19.48; Pred.No. 0.41; Indels 207; Gaps 29;

Matches 123; Conservative 90; Mismatches 213;

QY 5 GDNRRQCCGLADYLTSAKFLYLGHSLSTWGRMWHFAVSVFLVE----- 52

DB 17 GDYR-----GGLTVYVMVAFMAACGGLLGYDNGVTGGVSLFAFEKFFPDVNAKKQ 70

QY 53 -----LYGNSLLLTATYGLVGLSVLGAIGTDWVDKNAKLKVAO----- 93

DB 71 EVHEDSPYCYDNKQLQFLVSSFLAG---LVSLFASWITRWGRKVTWIGGAPFVAG 127

QY 94 --TSLVVQNVSVILCGIILM-----MVFLKHLLTMYHGWLTSYVILITIA-N 141

DB 128 GLVNAFAQDMWMLVGRVLGFGVLGSGVQVPPVLSFVAFPSRGMNLNIGYQLFVITIGIL 187

QY 142 IANLASTATAITIQDMVIVVAVAGEDSKLANMNATIRIDQLTNILAPNAV-QQIMTFGS 200

DB 188 IAGLVNAV-----RDW-----ENGWRLS-----LGPAAAPGAILFLGS 221

QY 201 PVIGCGFISGNLVSMCVYVLLWKVYOKTPALAVKAGLKEBETELKOLNHLKDTREKPL 260

DB 222 LVL-----PSPNPLVEKGTKEGREVLQ-----KL 247

QY 261 EGTLMGVKDSNIHELEHQEPTCASQMAEPRTFRDGMVSYNQ---PVFLAGMGLAFL 317

DB 248 CGTSEVDAEFADI-----VAAVEIARPI-TWQSWASLFTTRYMQLLTSFVIQF- 296

QY 318 YMTVLGFDICI-----TTGVAYTQGLSGSILSLMGASAITGIMGTVAFT-WLRRKC 367

DB 297 FOQFTGINAIEFVVPVLFSSLSGANSAL---LNTVVGVAVNVGSTLIAMFSDKEGRRF 353

QY 368 GLVPTGLISGLAQLSCLILCVISVFMPSG-PLDLSV-SPEDIRSHFIQES----- 417

DB 354 LIEGGIOQCCLAWLTGGVVLAEFAKYGTDPKAVASGILAVICIFISGFAMWGPWG 413

QY 418 IPTPKIPEITTEIYMSNGSNSANIVPESVPIISVLSLLFAGVIAAIGLWFSFLT 477

DB 414 LIPSEI-----FILETRP-----AGTAVAVVGNFLPSFVIG 444

QY 478 QLLQENVIESERGIINGVQNMVYLLDLHF-----IMVILAPNPAFGLLVLSVFA 532

DB 445 QAFVSMLCAMEYGV-----FLFFAGMLVIMVLC-----IFLLPETKGVP 484

QY 533 MGHIMYFRAQNTLGNKLFACGPDAGEVREKQ 565

DB 485 IERVOAL-YARHWFNRVM--GPAAAEVIAEDE 514

RESULT 4

YFIG_BACSU STANDARD; PRT; 482 AA.

AC P54723;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical metabolite transport protein yfig.

GN YFIG.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI_TaxID=1423;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=96262713; PubMed=8704981;

RA Yamanoto H., Uchiyama S., Fajar A.N., Ogasawara N., Sekiguchi J.;

RT "Determination of a 12 kb nucleotide sequence around the 76 degrees

RL Microbiology 142:1417-1421(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,

RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,

RA Borriess R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,

RA Chou S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,

RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,

RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,

RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,

RA Guiseppe G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holtsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,

RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,

RA Kobayashi Y., Koeter P., Koningsstein G., Krogh S., Kumano M.,

RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,

RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,

RA Noone D., Mellado R.P., Mizuno M., Moesti D., Nakai S., Noback M.,

RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,

RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,

RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seter S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandendool M., Vannier F., Vassarotti A.,
RA Viari A., Wambut R., Wedler H., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis".
RL Nature 390:249-256(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D50543; BAA09111.1; -;
CC PIR; B69803; B69803.
CC Subtilist; BG11854; yfig.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC InterPro; IPR005829; Sug_transporter.
CC InterPro; IPR003663; Sugar_transp.
CC Pfam; PF00083; sugar_tr; 1.
CC PRINTS; PR00171; SUGTRNSPORT.
CC TIGRfam; TIGR00879; SP; 1.
CC PROSITE; PS00216; SUGAR_TRANSPORT_1; FALSE_NEG.
CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
CC KW Hypothetical protein; Transmembrane; Complete proteome.
FT DOMAIN 1 29
FT TRANSMEM 30 50
FT DOMAIN 51 59
FT TRANSMEM 60 80
FT DOMAIN 81 92
FT TRANSMEM 93 113
FT DOMAIN 114 120
FT TRANSMEM 121 141
FT DOMAIN 142 155
FT TRANSMEM 156 176
FT DOMAIN 177 184
FT TRANSMEM 185 205
FT DOMAIN 206 263
FT TRANSMEM 264 284
FT DOMAIN 285 301
FT TRANSMEM 302 322
FT DOMAIN 323 331
FT TRANSMEM 332 352
FT TRANSMEM 353 373
FT DOMAIN 374 400
FT TRANSMEM 401 421
FT DOMAIN 422 423
FT TRANSMEM 424 444
FT DOMAIN 445 482
SQ SEQUENCE 482 AA; 52756 MW; 90851C4F4C48EE01 CRC64;
Query Match 3.9%; Score 113.5; DB 1; Length 482;
Best Local Similarity 19.7%; Pred. No. 0.83;
Matches 77; Conservative 73; Mismatches 154; Indels 87; Gaps 19;
Qy 34 LSTGDRMHWFAVSVFLVEL---YGNLLITAVVGLVAGSVLV---LGAIGCDWV-DK 85
Db 27 VSTGGLLFGYDTGVINGALPFMTAGQNLPTPTGVLVASSLLGAAFGAFMFGRLSDR 86
Qy 86 NARLKVAQTSVLVQNVSVILCGIIMVFLHKLHLLTMVHGWVLTSCVILIITIANL 145
Db 87 HGRRK-----TILYLALLFIAATL-----GCTSPNASVMIAFRFLGL 125
Qy 146 ASTATAITIQEDWIVVAGEDRSKLANNATIRIDOLTNILAPWAVQINTFSPVIGC 205
Db 126 AVGCASVTPTFLAISPARRGRIVTONELMIVIGQLLAYTFNAIIGSTWGESANV--- 182
Qy 206 GFISGMN--LVSMCEVYVLLW---KVYQKTPA-LAVKAGLKEEBTELKQLNLHDKTEPKP 259
Db 183 -----WYMLVIATLPVAVLWFGMLVIVPESPRWLAAGRMDALRVLRQ--IREDSQAQ- 234
Qy 260 LEGTHLMGVKDSNIHELEHEQEPTCASQMAEPPTFRDGVWVSYNQPVFLAGMGLAFLYM 319
Db 235 -----QETKEIKHAIEGTAKK---AGFHDFOEPMI---RRILFIGIA-IVQ 275
Qy 320 TVLGFDCI-----TTGVAYTQGLSGSTLSILMGASAITGIMGTVAFTWL---RRK 366
Db 276 QITGVNSIMYGYTEILREAGFQTEAALIGNI-----ANGVISVIAVIFGVLWLGKVRRR 329
Qy 367 CGLVRTGLISGLAQSLCLILCVISVFMFGSP 397
Db 330 PMLI-IGQIGTMTAL--LLIGILSVLEGT 357
RESULT 5
GTR2 RAT
ID GTR2 RAT STANDARD; PRT; 522 AA.
AC P12336;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 2
DE (Glucose transporter type 2, liver).
GN SLC2A2 OR GLUT2 OR GLUT-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89003066; PubMed=3048704;
RA Thorens B., Sarkar H.K., Kaback H.R., Lodish H.F.;
RT "Cloning and functional expression in bacteria of a novel glucose
RT transporter present in liver, intestine, kidney, and beta-pancreatic
RT islet cells".
RL Cell 55:281-290(1988).
CC -!- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOFORM LIKELY
CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
CC MEMBRANE OF HEPATOCYTES AND IS RESPONSIBLE FOR UPTAKE OF GLUCOSE
CC BY THE BETA CELLS; MAY COMPRISE PART OF THE GLUCOSE-SENSING
CC MECHANISM OF THE BETA CELL. MAY ALSO PARTICIPATE WITH THE
CC NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCELLULAR TRANSPORT OF
CC GLUCOSE IN THE SMALL INTESTINE AND KIDNEY.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: PRESENT IN LIVER, INTESTINE, KIDNEY AND
CC BETA-PANCREATIC ISLET CELLS.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC TRANSPORTER SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J03145; AAA41298.1; -;
CC PIR; A31556; A31556.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub_transporter.
CC InterPro; IPR005829; Sug_transporter.
CC InterPro; IPR003663; Sugar_transp.
CC Pfam; PF00083; sugar_tr; 1.

```

DR PRINTS: PR00171; SUGRTRANSPO.
DR TIGR00879; SP; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transport; Transport; Glycoprotein;
KW Multigene family.
DR DOMAIN 1 10 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 11 31 1 (POTENTIAL).
FT TRANSMEM 32 96 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 97 117 2 (POTENTIAL).
FT TRANSMEM 118 125 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 126 146 3 (POTENTIAL).
FT TRANSMEM 147 156 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 157 177 4 (POTENTIAL).
FT TRANSMEM 178 185 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 186 206 5 (POTENTIAL).
FT TRANSMEM 207 215 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 216 236 6 (POTENTIAL).
FT TRANSMEM 237 301 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 302 322 7 (POTENTIAL).
FT TRANSMEM 323 337 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 338 358 8 (POTENTIAL).
FT TRANSMEM 359 365 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 366 386 9 (POTENTIAL).
FT TRANSMEM 387 401 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 402 422 10 (POTENTIAL).
FT TRANSMEM 423 431 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 432 452 11 (POTENTIAL).
FT TRANSMEM 453 459 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 460 480 12 (POTENTIAL).
FT TRANSMEM 481 522 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 62 62 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 522 AA; 57085 MW; 075AB81B56CF33F7 CRC64;

Query Match 3.9%; Score 113.5; DB 1; Length 522;
Best Local Similarity 21.7%; Pred. No. 0.91;
Matches 102; Conservative 65; Mismatches 164; Indels 139; Gaps 26;

QY 52 ELYGNLLLTAVGLVAGSVL--VLGAIIGDWV-DKNARLKVAQTSLVVQNVSVILGCI 108
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 83 ETEGSAHIVTMLSLSVSPSPVAGVGVASFFGGLGDKLGRK----AMLAAN-SLSLTGA 137
QY 109 ILM--NVFLHKHLLT-----MYHGWV--LTSVILLIITIANIANLASTATAITQRD 157
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 138 LIMGCSKFGPAHALIAGRSVSGLYCLISGLVPMY-----IGEIAPT----- 181
QY 158 WIVVAGEDRSKLANMNATIRIDQLTNILAPMAVGQIMTFGSPVIGCGFISG-----WNL 213
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 182 -----LRGALGTLHQLALVTGILI-----SQIAGLSFILGNQDYWHI 218
QY 214 -----VSMCEVVLWKVYQKTPALAVKAGLKEEETELKQLNLHKDTEPKPLEGTHMG 267
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 219 LLGLSAPVALLQCLLL--LFCPSPRYLYLNLEEEVRAKKS-----KRLRGTEDI- 267
QY 268 VKDSNTHLEHEOEPTCSQMAQEPFRFDGWSVYNQPVFLAGM-GLAPLYMTVLGFDG 326
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 268 TKD--INEMKEKEASTEQKSVIQLFTD---PNYRQPIVVALMLHLAQFSGINGIFY 322
QY 327 ITTCYAVTQGLSGSIILMGASAITGIMGTVAFTWLRRCRGIVRTGLISG-LAQSLCLI 385
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 323 YSTISITQTAGISQPVYATI-GVGAINMIF-TAVSVLLVERAGR-RTLFLAGMIGMFFCAV 379-
QY 386 LCVISV-----FMPGSPDLDSVPEDRSRFSIQESTTPKIPBITTEIYWSNG 435
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 380 FMSLGLVLLDKFTWMSVYSMTAIFLFSVFE-----ICPGPIPMFMVAEFFSQ 428
QY 436 SNSANIVPETSPPSVPIIS-----VSLLPAGVI 463
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 429 -----PRATALAAAFSNWVCNFIALCFQIADFLGYPVFLFAGV 471

RESULT 6
YBQ_BACSU

```

```

ID YBQ_BACSU STANDARD; PRT; 435 AA.
AC E37489;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical transport protein ybq.
GN YBQ.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96051385; PubMed=7584024;
RA Ogasawara N., Nakai S., Yoshikawa H.;
RT "Systematic sequencing of the 180 kilobase region of the Bacillus
RT subtilis chromosome containing the replication origin.";
RL DNA Res. 1:1-14(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Chou S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.F.,
RA Entlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Paro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescan E., Fujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra F., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: BELONGS TO THE PHTHALATE PERMEASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D26185; BAA05188.1; -
CC EMBL; Z99124; CAB16094.1; -
CC PIR; S65982; S65982.
CC Subtilist; BG10016; YBQ.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Subtransporter.
CC Pfam; PF00083; sugar_tr; 1.
KW Hypothetical protein; Transmembrane; Transport; Complete proteome.
FT TRANSMEM 26 46 POTENTIAL.

```


QY 169 KLANNNATIR-----IDLNTNLA-----PMAVGQIMTSGSVICG 206
Db 55 QEAFLRQWQRLAQLHPQAFSLPWLQRITRNLRDLWRSHRRPLS-GEAADLA----- 106
QY 207 FISGNLVSMC-----VEYVLLWKVYQKTPALAVKAGLKEEETELKQLNLRKDTKPK 259
Db 107 -----TAMADPSPSPAEQAL--QVEEERAALEIMGALPNDSEI--LLYYREGORS 155
QY 260 LEGTHMGVQDSNIHELEHEQBPTCASQMAEPPFTFRDQ-----WVSYYNQPVPLAGMG 313
Db 156 QQVASLLGLSDAAVRKLSRABATVRNELLQRFDFARGSPGVAFATVTAATMLAAGP 215
QY 314 LAFVMTVLGDCI-----TTGVAYTQGLSGSILSILMGAS-----AITGIMGT 357
Db 216 TASAAIALGGIGSLGGVKGASGLSGSALTSGSAAAGLSVLLGMPMAIALLAITGVLT 275
QY 358 -----VAFTWLRKCLVR--TGLISGLAQLSC-----LILCVISVFM 393
Db 276 TYMSGAYLIRFATTAREAAAIINGFTRLSTLTAALTTCGAPLLRLRAFAPKWLALCVLAGM 335

RESULT 9

SL55_RAT STANDARD; PRT; 618 AA.
AC Q63008;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sodium/iodide cotransporter (Na(+)/I(-) cotransporter) (Sodium-iodide symporter) (Na+/I-symporter).
DE SLC5A5 OR NIS.
GN Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thyroid;
RX MEDLINE=96158880; PubMed=8559252;
RA Dai G., Levy O., Carrasco N.;
RT "Cloning and characterization of the thyroid iodide transporter.";
RL Nature 379:458-460(1996).
CC -!- FUNCTION: MEDIATES IODIDE UPTAKE IN THE THYROID GLAND.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE SODIUM: Solute symporter family (SSF).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U60282; AAB03338.1; -;
DR PIR; S68513;
DR InterPro; IPR001734; Na/solut_symport.
DR Pfam; PF00474; SSF; 1.
DR TIGRFAMs; TIGR00813; sss; 1.
DR PROSITE; PS00456; NA SOLUT_SYMP_1; 1.
DR PROSITE; PS00457; NA SOLUT_SYMP_2; FALSE_NEG.
DR PROSITE; PS0283; NA SOLUT_SYMP_3; 1.
KW Transport; Transmembrane; Sodium transport; Symport; Glycoprotein; Phosphorylation.
KW DOMAIN 1 16 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 17 37 POTENTIAL.
FT DOMAIN 38 53 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 54 74 POTENTIAL.
FT DOMAIN 75 88 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 89 109 POTENTIAL.
FT DOMAIN 110 136 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 137 157 POTENTIAL.
FT DOMAIN 158 163 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 164 184 POTENTIAL.
FT DOMAIN 185 186 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 187 207 POTENTIAL.
FT DOMAIN 208 241 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 242 262 POTENTIAL.
FT DOMAIN 263 286 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 287 307 POTENTIAL.
FT DOMAIN 308 326 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 327 347 POTENTIAL.
FT DOMAIN 348 391 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 392 412 POTENTIAL.
FT DOMAIN 413 416 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 417 437 POTENTIAL.
FT DOMAIN 438 444 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 445 465 POTENTIAL.
FT DOMAIN 466 520 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 521 541 POTENTIAL.
FT DOMAIN 542 618 CYTOPLASMIC (POTENTIAL).
FT MOD RES 551 551 PHOSPHORYLATION (BY PKA) (POTENTIAL).
FT CARBOHYD 485 485 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 618 AA; 65196 MW; 91EFAA6752B4F4B2 CRC64;
Query Match 3.7%; Score 108; DB 1; Length 618;
Best Local Similarity 19.1%; Pred. No. 2.8;
Matches 103; Conservative 71; Mismatches 160; Indels 204; Gaps 27;
QY 14 CGSLADYLTAKFLLYGLHSLTGWDRMWFHFAVSFLVELYGNLSLLLTAVYGLVVGSVL 73
Db 132 CGTL-QYLVAT--MLYTG-----IVYAPALILNQVTGLDIWASLL 169
QY 74 VLGAIGDWDKQARLKAQOTSLVQNVSVILCGIILMVFLHKLHLLTYHGVV-LTSC 132
Db 170 STGIICLYTVGGMKAVVWD--VFQVVVNLVGF-----WVILARG 209
QY 133 YILIITIANLASTATAITIQ-----RDWIVVWVAG----- 164
Db 210 VILGGPRNVLSLAQNSRLNLMDFDPDRSRYTFTFIVGGTLVWLSMYGVNQAVQRY 269
QY 165 -----EDRSKLANNNATIRIDQLTNLAPMAVGQIMTSGSVICGFSIGWNLVSMC 219
Db 270 VACHTEGKAKLAL-----VNQL-----GLFLIVASAAC-CG----- 300
QY 220 VVLLMKVYQK--TPALAVKAGLKEEETELKQLNLRKDTKPK-----LEGTH 264
Db 301 --IVMFYKDCDPLLTGRISAPDQYMLPLVDLDFEDLPVGLFLACAYSGTLSTASTS 358
QY 265 LMGVKDSNIHELEHEQBPTCASQMAEPPFTFRDQWVSYYNQPVPLAGMGLAFVMTVLG 324
Db 359 INAMAATVEDLIKPRMPGLAPR-----KLVPIS-KGLSFIY---GS 396
QY 325 DCITGVAYTQGLSGSILSILMGASAITG--INGTV-----AFT--WLRKCGL--VR 371
Db 397 ACLTV-----AALSSLLGGGVQLGSGFTVMGVISGFTLGAFTLMLLPACNTPGV 446
QY 372 TGLISGLAQLSCLILCVISVFMGSPDLVSFPEDIRSRFIGESITPKIPITTEIY 431
Db 447 SGLAAGLA-VSLWVAVGATLYPPGE-----QTMGVLPISAACTNDV 488
QY 432 M---SNGSNSANIVPETSPE-----SVPIISVSLIFAGVIAAARIGLWSFDLTVTQL 480
Db 489 LLGPPGATNASNGIPSSGMDTGRPALADTFVAISLYLYGALG-----TLTWTML 536
RESULT 10
YD19_METJA STANDARD; PRT; 492 AA.
AC Q58715;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical sodium-dependent transporter MJ1319.
GN MJ1319.


```
CC (Probable).
CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. SIMILAR TO
CC P. SYRINGAE SYRD.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; U00008; AA16403.1; -
CC EMBL; AE000310; AAC75271.1; -
CC EMBL; D90850; BAA15994.1; -
CC PIR; A64991; A64991.
CC EcoGene; EG12070; YoI.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR001140; ABC_TW_transpt.
CC InterPro; IPR003439; ABC_TRANSPORTER.
CC InterPro; IPR005898; Cyc pep transpt.
CC Pfam; PF00664; ABC_membrane; 1.
CC Pfam; PF00005; ABC_tran; 1.
CC ProDom; PD000006; ABC_transporter; 1.
CC SMART; SM00382; AAA; 1.
CC TIGRFAMs; TIGR01194; cyc pep trnspt; 1.
CC PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
CC PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
CC Hypothetical protein; ATP-binding; Transport; Transmembrane;
KW Inner membrane; Complete proteome.
KW TRANSMEM 16 36 POTENTIAL.
FT TRANSMEM 48 68 POTENTIAL.
FT TRANSMEM 125 145 POTENTIAL.
FT TRANSMEM 149 169 POTENTIAL.
FT TRANSMEM 240 260 POTENTIAL.
FT TRANSMEM 270 290 POTENTIAL.
FT NP_BIND 356 363 ATP (POTENTIAL).
SQ SEQUENCE 547 AA; 61552 MW; 11523935EDFD9EAB CRC64;
Query Match 3.6%; Score 106; DB 1; Length 547;
Best Local Similarity 20.6%; Pred. No. 3.3; Indels 116; Gaps 16;
Matches 78; Conservative 55; Mismatches 130;
QY 37 WGRMRHFAVSFLVELYNSLLTAVYGLVVGSLVVLGAIIGDWDVKNARL-KVAQTS 95
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
8 WRQYRWPF-----ISVWALSASALGIGLI-----AFINQRLIETADTS 47
QY 96 LVV--QNVSVILGIIIMVFLKH-ELLTWHGV--LTSYILITITIANIANLASTAT 150
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
48 LLVLPFLGLL---LLMAVTIGSQLATTTLGHFVYRLRSEFIKRIIDTHVERIEQLGS 104
QY 151 AITTIQDWIVVVGEDRSKLANMNATIRIDQLTNILAPMAVGQIMTFGSPVCGGFI 210
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
105 A-----SLLAGLTSVDRNITIAFVRUPELVQGLITLIGSAAY-LMWLSG 147
QY 211 -----WNLVSMCEVYLLMKVYQKTPALAVKAGLKEETEL-----KOLNL 251
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
148 KMLLVTAIWMAITWGPFVLVARYKH-----WATLRETDKLYTDFQVLEGRKELTL 201
QY 252 HKDTEPKPLEGTHLMGVKDSNIHELEHQEPTCASQMAEPRTRFDGQVSVYNOVFVLAG 311.
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
202 NRERAEEYFNLLYPDAQEYRHHIR-----ADTFHLSAVNVW-----SNIMMLGA 246
QY 312 MGLAFYMTVLGFDCTITGYATQGLSGSIILMGASAITGIMGTVAFTWLRKCKGLVR 371
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
247 IGLVFWMANSJGW-----ADTNVAATVSLTLL-----FLR 276
QY 372 TGLISGLAQLSCLILCVIS 390
Db | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
277 TPLLSAVGALPTLLTAQVA 295
```

```
GTRS RAT
ID GTRS RAT STANDARD; PRT; 502 AA.
AC P43427;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 5
DE Glucose transporter type 5, small intestine (Fructose transporter).
GN SLC2A5 OR GLUT5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Jejunum;
RX MEDLINE=93325725; PubMed=8333543;
RA Rand E.B., Depaoli A.M., Davidson N.O., Bell G.I., Burant C.F.;
RT "Sequence, tissue distribution, and functional characterization of
RT the rat fructose transporter GLUT5.";
RL Am. J. Physiol. 264:G1169-G1176(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Jejunum;
RX MEDLINE=9408761; PubMed=8404647;
RA Inukai K., Asano T., Katagiri H., Ishihara H., Anai M.,
RA Fukushima Y., Tsukuda K., Kikuchi M., Yazaki Y., Oka Y.;
RT "Cloning and increased expression with fructose feeding of rat jejunal
RT GLUT5.";
RL Endocrinology 133:2009-2014(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Intestine;
RA Kasahara T., Kasahara M.;
RA Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CYTOCHALASIN B-SENSITIVE CARRIER. SEEMS TO FUNCTION
CC PRIMARILY AS A FRUCTOSE TRANSPORTER.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN SMALL INTESTINE, AND AT MUCH
CC LOWER LEVELS IN KIDNEY, SKELETAL MUSCLE, AND ADIPOSE TISSUE.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC TRANSPORTER SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; L05195; AAA02627.1; -
CC EMBL; D13871; BAA02983.1; -
CC EMBL; D28562; BAA05912.1; -
CC PIR; I53268; I53268.
CC InterPro; IPR007114; MFS.
CC InterPro; IPR005828; Sub transporter.
CC InterPro; IPR005829; Sug_transporter.
CC InterPro; IPR003663; Sug_transpt.
CC Pfam; PF00083; sugar_tr; 1.
CC PRINTS; PR00171; SUGRTRANSPORT.
CC TIGRFAMs; TIGR00879; SP; 1.
CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
CC Transmembrane; Sugar transport; Transport; Glycoprotein;
KW Multigene family.
KW DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 12 32 1 (POTENTIAL).
FT DOMAIN 33 71 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 72 92 2 (POTENTIAL).
FT DOMAIN 93 99 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 100 120 3 (POTENTIAL).
FT DOMAIN 121 131 EXTRACELLULAR (POTENTIAL).
```



```
FT TRANSMEM 132 152 4 (POTENTIAL).
FT DOMAIN 153 160 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 161 181 5 (POTENTIAL).
FT DOMAIN 182 191 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 192 212 6 (POTENTIAL).
FT DOMAIN 213 276 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 277 297 7 (POTENTIAL).
FT DOMAIN 298 314 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 315 335 8 (POTENTIAL).
FT DOMAIN 336 342 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 343 363 9 (POTENTIAL).
FT DOMAIN 364 367 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 368 388 10 (POTENTIAL).
FT DOMAIN 389 411 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 412 432 11 (POTENTIAL).
FT DOMAIN 433 438 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 439 459 12 (POTENTIAL).
FT DOMAIN 460 502 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 164 164 V -> A (IN REF. 1).
FT CONFLICT 285 285 A -> T (IN REF. 2).
FT CONFLICT 452 452 T -> S (IN REF. 2).
FT CONFLICT 479 479 K -> N (IN REF. 1).
SQ SEQUENCE 502 AA; 55543 MW; 871FEDEFA1E7894D CRC64;

Query Match
Best Local Similarity 3.6%; Score 104.5; DB 1; Length 502;
Matches 121; Conservative 88; Mismatches 211; Indels 185; Gaps 31;

QY 15 GSLADYLTSAKFLYLGHSLSTWGRMHFPAVSVFLVELYXGN-----SLLLTAV 63
DB 10 GKLTVLALATFLAAGSSFOYGVNVAVNSPSEFMQFQYNDVYDRNKENIESFTLLTLL 69
QY 64 YGLVVA-----GSVLVGAIGLDWVDKARL-----KVAQT--SLV 97
DB 70 WSLTVSPFPFGFPIGLSLMV-GFLVNNLGRKALLFNFIIPAILMGCSKIAKSPFII 128
QY 98 VQNVSVILC-GIILMMVFLHKLHLLTMVHGWLVTSCVILLIITIA-NIANLASTATAITIQ 155
DB 129 ASRLLVGICAGISNVVPMVGLGAPKNGALGVVLPQLEITVGLVLAQLFGLRSVLASE 188
QY 156 RDMIVVAGEDRSKLANMNATIRIDQLTNILAPMAVGQIMTF-GSPVIGCGFISGNLV 214
DB 189 EGNPILIG-----LTGVPAQLQLLLPFPFESP----- 216
QY 215 SMCVEVULLMKVQKTPALAVKACGLKEEETELQNLNKHDKTEPKPLEGTHLMGVKDSNIH 274
DB 217 -----RYLLIQK--KNESAIEKA-----LQTLRGWKDV-----DMEME 247
QY 275 ELEHEQBPCTCASQMAEPFRIFRD---GWVSYYNOPVFLAGMGLAFLYMTVLGFDCTITGY 331
DB 248 EIRKEDEAEKAGFISVWKLFRMQSLRW-QLISTIVLMAGQQLS-----GVNAI--Y 296
QY 332 AYTOGL-----SGSILSILMGASAITGIMGTAVF---TWLRRKGLVTRTGLISGLAQ 380
DB 297 YYADQIVLSAGKSNVDVQVYTAGTVGAVNVFMTVTVFVVELWGRNL-----LLIGFS- 349
QY 381 LSCILCVISVFMGSPGLDLSVSPFEDIRS--RFIQGESITPKIPEI-TTEIWMNGSN 437
DB 350 -TCLTACIVLTVALALQNTISWMPYVSVICVIVVIVGHVGPSPFIPALFITEIFLOSSRP 408
QY 438 SANIVPETSPESPPIIS---VSLLFAGVIAARIGLSFDTLTQQLQENVIESERGIING 494
DB 409 SAYMI-----GGSVHWSNFIIVGLIFPFI---QVGLGYPFIIIFAI----- 447
QY 495 VQNSMNYLLDLHLHFIMVILAPNEAFGLLVLIISVFMAGHIMYFRFAQNTLGNKLFACG 554
DB 448 -----CLLTIVIFMVV-----PETKGR-----TFVEINQI---FAKK--NKVSDVY 484
QY 555 PDAKE 559
DB 485 PEKEE 489
```

```
RESULT 13
A2AC_CAVPO STANDARD; PRT; 455 AA.
ID A2AC_CAVPO
AC Q60476;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Alpha-2C adrenergic receptor (Alpha-2C adrenoceptor).
GN ADRA2C.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hartley;
RA MEDLINE=96152573; PubMed=8573196;
RX Svensson S.P., Bailey T.J., Porter A.C., Richman J.G., Regan J.W.;
RT "Heterologous expression of the cloned guinea pig alpha 2A, alpha 2B,
RT and alpha 2C adrenoceptor subtypes. Radioligand binding and
RT functional coupling to a CAMP-responsive reporter gene.";
RL Biochem. Pharmacol. 51:291-300(1996).
CC -!- FUNCTION: ALPHA-2 ADRENERGIC RECEPTORS MEDIATE THE CATECHOLAMINE-
CC INDUCED INHIBITION OF ADENYLATE CYCLASE THROUGH THE ACTION OF G
CC PROTEINS.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U25724; AAA67076.1; -.
CC HSSP; P29274; LMMH.
CC InterPro; IPR000276; GPCR_Rhodopsn.
CC Pfam; PF00001; 7tm1; 1.
CC PRINTS; PR00237; GPCRHHODOPS.
CC PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
CC PROSITE; PS00262; G-PROTEIN RECEPTOR FL2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Multigene family; Phosphorylation
KW DOMAIN 1 47 EXTRACELLULAR (POTENTIAL).
KW TRANSMEM 48 72 1 (POTENTIAL).
KW DOMAIN 73 84 CYTOPLASMIC (POTENTIAL).
KW TRANSMEM 85 110 2 (POTENTIAL).
KW DOMAIN 111 120 EXTRACELLULAR (POTENTIAL).
KW TRANSMEM 121 143 3 (POTENTIAL).
KW DOMAIN 144 164 CYTOPLASMIC (POTENTIAL).
KW TRANSMEM 165 187 4 (POTENTIAL).
KW DOMAIN 188 203 EXTRACELLULAR (POTENTIAL).
KW TRANSMEM 204 227 5 (POTENTIAL).
KW DOMAIN 228 237 6 (POTENTIAL).
KW TRANSMEM 237 400 EXTRACELLULAR (POTENTIAL).
KW DOMAIN 401 413 7 (POTENTIAL).
KW TRANSMEM 414 434 8 (POTENTIAL).
KW DOMAIN 435 455 CYTOPLASMIC (POTENTIAL).
KW FT CARBOHYD 17 17 N-LINKED (GLCNAC. .) (POTENTIAL).
KW FT CARBOHYD 29 29 N-LINKED (GLCNAC. .) (POTENTIAL).
KW FT DISULFID 120 198 BY SIMILARITY.
KW SQ SEQUENCE 455 AA; 68657D247FF8A3F4 CRC64;

Query Match
Best Local Similarity 3.5%; Score 103.5; DB 1; Length 455;
Matches 54; Conservative 36; Mismatches 80; Indels 49; Gaps 11;

QY 60 LTAVYGLV-----VAGSVLVGLAIGDWVDKARLKVQATSLVVQNVSVILCGIILMMVFL 115
DB 50 LAADVGLIVFTVGVNVLVIAVL-----TSRALRAPQNLFLVSLASADILVATLVMPFS 104
```


QY 116 HKHELLT-MYHGWLTSCVI---LIITIANIANLASTATAITQDRIWVWVAGEDRSKLA 171
Db 105 LANELLMAYWYFQVWGVYALDVLCTSSIVHLC---AISLDYWSVTQAVE-----Y 155
QY 172 NNATIRIDQLTNILAPWVQIMTF-----GSPVIGGPF-ISGNWLVSMCV- 218
Db 156 NLKTRPRVK--ATIVAVWLISAIISFPLVSVFYRQPDGAAYPRGNDLDETWYILSSCIG 213
QY 219 -----EYVLMLKVQY-KTPALAVKAGLKEEETE 245
Db 214 SFFAPCLIMGLVYARIYVAKLRTYTLSEKRGPAPEGE 252
RESULT 14
GTR2_CHICK
ID GTR2_CHICK STANDARD; PRT; 533 AA.
AC Q90592;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 2
DE (Glucose transporter type 2, liver).
GN SLC2A2 OR GLUT2 OR GLUT-2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94213485; PubMed=8161202;
RA Wang M.Y., Tsai M.Y., Wang C.;
RT "Identification of chicken liver glucose transporter.";
RL Arch. Biochem. Biophys. 310:172-179(1994).
CC -!- FUNCTION: FACILITATIVE GLUCOSE TRANSPORTER. THIS ISOPFORM LIKELY
CC MEDIATES THE BIDIRECTIONAL TRANSFER OF GLUCOSE ACROSS THE PLASMA
CC MEMBRANE OF HEPATOCYTES AND IS RESPONSIBLE FOR UPTAKE OF GLUCOSE
CC BY THE BETA CELLS; MAY COMPRISE PART OF THE GLUCOSE-SENSING
CC MECHANISM OF THE BETA CELL. MAY ALSO PARTICIPATE WITH THE
CC NA(+)/GLUCOSE COTRANSPORTER IN THE TRANSCYLLULAR TRANSPORT OF
CC GLUCOSE IN THE SMALL INTESTINE AND KIDNEY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
CC TRANSPORTER SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z22932; CAA80519.1; -;
DR PIR; S43230;
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR Pfam; PF00083; sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRFAMS; TIGR00879; SP; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00211; SUGAR_TRANSPORT_2; 1.
KW Transmembrane; Sugar transporter; Transport; Glycoprotein;
KW Multigene family.
FT DOMAIN 1 17 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 18 38 1 (POTENTIAL).
FT DOMAIN 39 110 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 111 131 2 (POTENTIAL).
FT DOMAIN 132 136 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 137 157 3 (POTENTIAL).

FT DOMAIN 158 163 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 164 184 4 (POTENTIAL).
FT DOMAIN 185 199 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 200 220 5 (POTENTIAL).
FT DOMAIN 221 229 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 230 250 6 (POTENTIAL).
FT DOMAIN 251 315 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 316 336 7 (POTENTIAL).
FT DOMAIN 337 350 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 351 371 8 (POTENTIAL).
FT DOMAIN 372 379 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 380 400 9 (POTENTIAL).
FT DOMAIN 401 413 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 414 434 10 (POTENTIAL).
FT DOMAIN 435 445 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 446 466 11 (POTENTIAL).
FT DOMAIN 467 471 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 472 492 12 (POTENTIAL).
FT DOMAIN 493 533 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 64 64 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 69 69 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 533 AA; 57699 MW; D06284CDC2779A3D CRC64;
Query Match 3.5%; Score 103.5; DB 1; Length 533;
Best Local Similarity 20.2%; Pred. No. 4.9;
Matches 89; Conservative 70; Mismatches 129; Indels 153; Gaps 23;
QY 42 WHFAVSVFLVELYGNLSLLTAVYGLVWAGSVLVGLGIIIGDWV-DKNARKVQAQTSVVQN 100
Db 109 WSLSVSMF-----AVGWSVSFT-----GWIGRLGRVK-----AMLVVN 144
QY 101 VSVILCGIILMVFHLKHHELLTWYHGVLTSCYILITIANIANLASTATAITQDRIWV 160
Db 145 VLSAGNLLMGLAKMGSHIL-IIAGRAITGLY-----CGLSSGLVP-----M 186
QY 161 VWAGEDRSKLANNNATIRIDQLTNILAPWVQIMTFGSPVIGCGFISG-----WNL--- 213
Db 187 YVSEVSPALTALRGALGTUHLQAIIVTGILI-----SQVLGLDFLGLNDELWPLLLG 235
QY 214 ---VSMCEVVELLWKVYQKTPALAVKAG-LKEBETELKQLNLHKDTPKPLEGTHLMGVK 269
Db 236 LSGVAALLQFFLLLLCPESPRIYIKLGRVEEAKSKLRLRGNC-----PMK----- 283
QY 270 DSNHLEHEQEPFCASQMAEPFTRDGMVSVYVNPVELA-----GMGLAFLYM 319
Db 284 --EIAEMEKEQEAASEKRVISIGOLFSS-----SKYQAVIVALMVQISQQSGINAFYYS 338
QY 320 T-----VLGPDCTITTYAYTQGLSGSILSLMGASA-----ITGIMG--- 356
Db 339 TNIFQAGVGPVYVATIGVGWNTVF-----TVISVFLVEKAGRRSLFLAGLMGLI 390
QY 357 -TVAFWLRKCGLVRTGLISGLAQSLCLLCVISMVPM-----PG----- 395
Db 391 SAVAMT-----VGLV-----LLSQFAMWSYVSMVAIFLFEVFEVGGPIPMFVIAELFSQ 442
QY 396 -SPDLISVSPFEDIRSFIOG 415
Db 443 PRPAATAVAGFCNACNFIVG 463
RESULT 15
HH2R_RAT
ID HH2R_RAT STANDARD; PRT; 358 AA.
AC P25102;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Histamine H2 receptor (H2R) (Gastric receptor 1).
GN HHR2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;

```

[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=92028890; PubMed=1930186;
RA Ruat M., Traflet E., Arrang J.-M., Leurs R., Schwartz J.-C.;
RT "Cloning and tissue expression of a rat histamine H2-receptor gene.";
RL Biochem. Biophys. Res. Commun. 179:1470-1478(1991).
CC -1- FUNCTION: THE H2 SUBCLASS OF HISTAMINE RECEPTORS MEDIATES GASTRIC
CC ACID SECRETION. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G
CC PROTEINS WHICH ACTIVATE ADENYLYL CYCLASE.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; S57565; BAB19935.1; -.
DR PIR; JQ1278; JQ1278.
DR HSP; P29274; IMM.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT DOMAIN 1 22 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 23 44 1 (POTENTIAL).
FT DOMAIN 45 57 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 58 81 2 (POTENTIAL).
FT DOMAIN 82 92 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 93 114 3 (POTENTIAL).
FT DOMAIN 115 134 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 135 159 4 (POTENTIAL).
FT DOMAIN 160 179 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 180 203 5 (POTENTIAL).
FT DOMAIN 204 233 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 234 257 6 (POTENTIAL).
FT DOMAIN 258 266 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 267 288 7 (POTENTIAL).
FT DOMAIN 289 358 CYTOPLASMIC (POTENTIAL).
FT SITE 98 98 ESSENTIAL FOR HISTAMINE BINDING (BY
FT SITE 185 185 SIMILARITY).
FT SITE 189 189 ESSENTIAL FOR TIOTIDINE BINDING AND
FT SITE 189 189 IMPLICATED IN HISTAMINE BINDING (BY
FT SITE 189 189 IMPLICATED IN HISTAMINE BINDING (BY
FT CARBOHYD 4 4 N-LINKED (GLCNAC... ) (POTENTIAL).
FT DISULFID 91 173 BY SIMILARITY.
SQ SEQUENCE 358 AA; 40253 MW; 4889F69B7B5D5DDC CRC64;

Query Match 3.5%; Score 103; DB 1; Length 358;
Best Local Similarity 19.8%; Pred. No. 3.3;
Matches 66; Conservative 53; Mismatches 112; Indels 102; Gaps 13;

QY 57 SLLLTAVYGLVAGSVLVGLGAIIGDWKDKARLKVAQTSILVQVNSVILCGIILMMVFLH 116
DB 22 SVVLTLLITLTIAGNVVCLIA-----VSLNRRLSLTNCFCIVSLAATDLLGLVLPFSA 76
QY 117 KHEL-LTMVHGWLTCYI---LIITIANI-----ASTATA 151
DB 77 IYQLSFETSPSGHVFNCNIYSLDVMLCTASILNLSLDRCYCAVTDPLRYPVLVTPVRVA 136
QY 152 IITQRDWIVVV-----AGEDRSKLANMNATIRRIIDLTNLAIPMAVGQIMTFGSPV 202
DB 137 ISLVFTWVITSLFSLSIHGWNRSNRGTGCGNDTFKCKVQVNEVYG--LVDGLVTFYLP- 193.
QY 203 ICGCFISGWNLVSMCVVEYVLLWKVYQKTPALAVKAGLKEETELKQLNLHKDTEPKPLEG 262

```

```

Db 194 -----LLIMCVYYRIKFIARE-----QAKRIN----- 216
QY 263 THLMGVKDSNIHELEHEQEPTCASQMAEPRTFDCGWVSYYNOPVFLAGMGLAFLYMTVL 322
DB 217 -HSSWKAATIR--EHKATVTLLAAVNG-----AFIICWFFPYFT-----AFVYRGLR 259
QY 323 GPDCTITGVAYTQGLSGSILSLMGASAITGIM 355
DB 260 GDAI-----NEAVEGIVLWLGYSANLNPIL 286

```

Search completed: November 5, 2003, 18:31:18
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 18:24:21 ; Search time 58 Seconds
(without alignments)
2540.483 Million cell updates/sec

Title: US-09-715-927-6

Perfect score: 2929

Sequence: 1 MTRAGDHNRRQCCGLADY.....ACGPDKEVRKENQANTSVV 571

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteria:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2929	100.0	571	Q9NP59	Q9np59 homo sapien
2	2921	99.7	571	Q8IVB2	Q8ivb2 homo sapien
3	2906	99.2	571	Q9NRL0	Q9nrl0 homo sapien
4	2618.5	89.4	570	Q9JHI9	Q9jhi9 mus musculus
5	2612.5	89.2	570	Q9JKP4	Q9jhp4 mus musculus
6	2600.5	88.8	570	Q923U9	Q923u9 rattus norv
7	1975.5	67.4	562	Q8JFW0	Q8jfw0 brachydanio
8	1971.5	67.3	562	Q919R3	Q919r3 brachydanio
9	1894.5	64.7	528	Q8AW23	Q8aw23 brachydanio
10	1513.5	51.7	436	Q8AW28	Q8aw28 brachydanio
11	1479.5	50.5	405	Q9Z1C9	Q9z1c9 rattus norv
12	1360.5	46.4	305	Q8BME5	Q8bme5 mus musculus
13	859	29.3	167	Q9NUS1	Q9nus1 homo sapien
14	612.5	20.9	524	O80905	O80905 arabidopsis
15	564	19.3	498	Q9L2C8	Q9l2c8 arabidopsis
16	509.5	17.4	551	O44607	O44607 caenorhabdi

17	498	17.0	485	5	Q8IA95	Q8ia95 caenorhabdi
18	471.5	16.1	440	5	Q9BKT8	Q9bkt8 caenorhabdi
19	371	12.7	81	11	Q8BUM5	Q8bum5 mus musculu
20	301	10.3	598	10	Q8W4E7	Q8w4e7 arabidopsis
21	209.5	7.2	677	5	O01891	O01891 caenorhabdi
22	187.5	6.4	482	10	O04629	O04629 arabidopsis
23	164	5.6	37	11	Q9JIM9	Q9jim9 mus musculu
24	157	5.4	496	16	Q8G5T4	Q8g5t4 bifidobacte
25	136.5	4.7	713	16	Q983A3	Q983a3 rhizobium l
26	135.5	4.6	412	16	Q8Y1P3	Q8y1p3 ralsstonia s
27	128	4.4	1287	5	Q9V6Z6	Q9v6z6 drosophila
28	123.5	4.2	409	16	Q97LU6	Q97lu6 clostridium
29	120.5	4.1	393	16	Q8G2C7	Q8g2c7 brucella su
30	120.5	4.1	763	16	Q8YFI5	Q8yfi5 brucella me
31	120	4.1	421	16	Q8G4V4	Q8g4v4 bifidobacte
32	120	4.1	481	5	Q9VNM1	Q9vnm1 drosophila
33	120	4.1	485	5	Q8S235	Q8sz35 drosophila
34	119.5	4.1	949	16	Q8FTZ6	Q8ftz6 corynebacte
35	119	4.1	537	16	Q8NLL2	Q8nll2 corynebacte
36	118	4.0	437	16	Q8YTP1	Q8ytp1 anabaena sp
37	117.5	4.0	783	10	O65383	O65383 arabidopsis
38	116	4.0	469	17	Q97VB9	Q97vb9 sulfolobus
39	115.5	3.9	418	16	Q9KG22	Q9kg22 bacillus ha
40	115.5	3.9	491	5	Q9V8N1	Q9v8n1 drosophila
41	115	3.9	408	16	Q97TP9	Q97tp9 clostridium
42	115	3.9	503	5	Q9XW19	Q9xw19 caenorhabdi
43	114.5	3.9	387	2	Q9KWU2	Q9kwu2 sphingomona
44	114.5	3.9	424	16	Q8NU27	Q8nu27 corynebacte
45	114	3.9	392	16	Q97QT3	Q97qt3 streptococc

ALIGNMENTS

RESULT 1

Q9NP59 ID Q9NP59 PRELIMINARY; PRT; 571 AA.
 AC Q9NP59;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DE 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Iron-regulated transporter IREG1 (Ferroporlin) (Hypothetical protein)
 DE (Solute carrier family 11 (Proton-coupled divalent metal ion transporters), member 3).
 DE IREG1 OR FPNI OR DKFZP586J0624.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20337919; PubMed=10882071;
 RA McKie A.T., Marciani P., Rolfs A., Brennan K., Wehr K., Barrow D.,
 RA Miret S., Bomford A., Peters T.J., Farzaneh F., Hediger M.A.,
 RA Hentze M.W., Simpson R.J.;
 RA "A novel duodenal iron-regulated transporter, IREG1, implicated in the basolateral transfer of iron to the circulation.";
 RL Mol. Cell 5:299-309(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Placenta;
 RA MEDLINE=20155474; PubMed=10693807;
 RA Donovan A., Brownlie A., Zhou Y., Shepard J., Pratt S.J., Moynihan J.,
 RA Paw B.H., Drejer A., Barut B., Zapata A., Law T.C., Brugnara C.,
 RA Lux S.E. IV, Pinkus G.S., Pinkus J.L., Kingsley P.D., Palis J.,
 RA Fleming M.D., Andrews N.C., Zon L.I.;
 RA "Positional Cloning of Zebrafish Ferroportin1 Identifies a Conserved Vertebrate Iron Exporter.";
 RL Nature 403:776-781(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Uterus;
 RA Duesterhoeft A., Lauber J., Mewes H.W., Gassenhuber J., Wiemann S.;

```

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF231121; AAF44330.1; -
DR EMBL; AF226614; AAF36697.1; -
DR EMBL; AL136944; CAB66878.1; -
DR EMBL; BC037733; AAH37733.1; -
DR EMBL; BC037733; AAH37733.1; -
DR Gnew; HGNC:10909; SLC11A3.
DR InterPro; IPR001993; Mitoch carrier.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Hypothetical protein.
SQ SEQUENCE 571 AA; 62542 MW; 54D6B5594C904959 CRC64;

Query Match 100.0%; Score 2929; DB 4; Length 571;
Best Local Similarity 100.0%; Pred. No. 1.3e-216;
Matches 571; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTRAGDHNRGCGGSLADYLTSAKFLLYLGHSLSTWGDMMHFAVSFVFLVLYGNSLL 60
DB 1 MTRAGDHNRGCGGSLADYLTSAKFLLYLGHSLSTWGDMMHFAVSFVFLVLYGNSLL 60
QY 61 TAVYGLVWAGSVLVGLGAIIGDWDKNARLKVAQTSLVVQNVSVILCGIILMMVFLKH 120
DB 61 TAVYGLVWAGSVLVGLGAIIGDWDKNARLKVAQTSLVVQNVSVILCGIILMMVFLKH 120
QY 121 LTMVHGWLTSYLIITITIANIANLASTATAITIQRDWIVVAGEDRSKLANNNATIR 180
DB 121 LTMVHGWLTSYLIITITIANIANLASTATAITIQRDWIVVAGEDRSKLANNNATIR 180
QY 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGNNLVSMCVYVLLWKYQKTPALAVKAG 240
DB 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGNNLVSMCVYVLLWKYQKTPALAVKAG 240
QY 241 EETELKQLNLHKDTEPKLEGTHLMGVKDSNIHELEHEOEPTCASQMAEPFTRDGV 300
DB 241 EETELKQLNLHKDTEPKLEGTHLMGVKDSNIHELEHEOEPTCASQMAEPFTRDGV 300
QY 301 SYYNQPVFLAGMGLAFLYMTVLGFCDCITTGAYATQGLSGSILSLMGASAITGIM 360
DB 301 SYYNQPVFLAGMGLAFLYMTVLGFCDCITTGAYATQGLSGSILSLMGASAITGIM 360
QY 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPFEDIRSRFTQGES 420
DB 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPFEDIRSRFTQGES 420
QY 421 TKIPEITTEIYMGNSNSANIVPETSPEVPIISVLLPAGVIAARIGLWSFDLTVTQL 480
DB 421 TKIPEITTEIYMGNSNSANIVPETSPEVPIISVLLPAGVIAARIGLWSFDLTVTQL 480
QY 481 QENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPEAFGLLVLSVSVFAMGHIM 540
DB 481 QENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPEAFGLLVLSVSVFAMGHIM 540
QY 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSV 571
DB 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSV 571

RESULT 2
Q81VB2 ID Q81VB2 PRELIMINARY; PRT; 571 AA.
AC Q81VB2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Solute carrier family 11 (Proton-coupled divalent metal ion transporters), member 3.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035893; AAH35893.1; -
DR EMBL; BC035893; AAH35893.1; -
SQ SEQUENCE 571 AA; 62482 MW; 45D6B5503C90495F CRC64;

Query Match 99.7%; Score 2921; DB 4; Length 571;
Best Local Similarity 99.8%; Pred. No. 5.5e-216;
Matches 570; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTRAGDHNRGCGGSLADYLTSAKFLLYLGHSLSTWGDMMHFAVSFVFLVLYGNSLL 60
DB 1 MTRAGDHNRGCGGSLADYLTSAKFLLYLGHSLSTWGDMMHFAVSFVFLVLYGNSLL 60
QY 61 TAVYGLVWAGSVLVGLGAIIGDWDKNARLKVAQTSLVVQNVSVILCGIILMMVFLKH 120
DB 61 TAVYGLVWAGSVLVGLGAIIGDWDKNARLKVAQTSLVVQNVSVILCGIILMMVFLKH 120
QY 121 LTMVHGWLTSYLIITITIANIANLASTATAITIQRDWIVVAGEDRSKLANNNATIR 180
DB 121 LTMVHGWLTSYLIITITIANIANLASTATAITIQRDWIVVAGEDRSKLANNNATIR 180
QY 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGNNLVSMCVYVLLWKYQKTPALAVKAG 240
DB 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGNNLVSMCVYVLLWKYQKTPALAVKAG 240
QY 241 EETELKQLNLHKDTEPKLEGTHLMGVKDSNIHELEHEOEPTCASQMAEPFTRDGV 300
DB 241 EETELKQLNLHKDTEPKLEGTHLMGVKDSNIHELEHEOEPTCASQMAEPFTRDGV 300
QY 301 SYYNQPVFLAGMGLAFLYMTVLGFCDCITTGAYATQGLSGSILSLMGASAITGIM 360
DB 301 SYYNQPVFLAGMGLAFLYMTVLGFCDCITTGAYATQGLSGSILSLMGASAITGIM 360
QY 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPFEDIRSRFTQGES 420
DB 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPFEDIRSRFTQGES 420
QY 421 TKIPEITTEIYMGNSNSANIVPETSPEVPIISVLLPAGVIAARIGLWSFDLTVTQL 480
DB 421 TKIPEITTEIYMGNSNSANIVPETSPEVPIISVLLPAGVIAARIGLWSFDLTVTQL 480
QY 481 QENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPEAFGLLVLSVSVFAMGHIM 540
DB 481 QENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPEAFGLLVLSVSVFAMGHIM 540
QY 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSV 571
DB 541 FAQNTLGNKLFACGPDPAKEVRKENQANTSV 571

RESULT 3
Q9NRL0 ID Q9NRL0 PRELIMINARY; PRT; 571 AA.
AC Q9NRL0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SLC11A3 iron transporter.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20347902; PubMed=10747949;
RA Abboud S., Haile D.J.;
RT "A Novel Mammalian Iron-regulated Protein Involved in Intracellular Iron Metabolism."
RT J. Biol. Chem. 275:19906-19912(2000).
RL J. Biol. Chem. 275:19906-19912(2000).
```

```

RN  SEQUENCE FROM N.A.
RP  MEDLINE=20291023; PubMed=10828623;
RA  Haile D.J.;
RT  "Assignment of slc11a3 to mouse chromosome 1 band 1B and SLC11A3 to
RL  human chromosome 2q32 by in situ hybridization.";
DR  EMBL; AF215636; AAF80986.1; -
DR  InterPro; IPR001993; Mitoch carrier.
DR  PROSITE; PS00215; MITOCH_CARRIER; 1.
SQ  SEQUENCE 571 AA; 62581 MW; 887078FFC4A4B921 CRC64;

Query Match      99.2%; Score 2906; DB 4; Length 571;
Best Local Similarity 99.5%; Pred. No. 7.8e-215;
Matches 568; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy  1 MTRAGDHNRRGCCGSLADYLSAKFLYLGHSLSTWGRMWHFAVSFLVELYGNLSLL 60
Db  1 MTRAGDHNRRGCCGSLADYLSAKFLYLGHSLSTWGRMWHFAVSFLVELYGNLSLL 60
Qy  61 TAVYGLVAGSVLVLGAILGDWVDKNARLKVAQTSLVVQNVSVILCGIILMMVFLHKL 120
Db  61 TAVYGLVAGSVLVLGAILGDWVDKNARLKVAQTSLVVQNVSVILCGIILMMVFLHKL 120
Qy  121 LTMVHGWLTSYLIITIANIANLASTATAITIQRDWIVVWAGEDRSKLANMATIRRI 180
Db  121 LTMVHGWLTSYLIITIANIANLASTATAITIQRDWIVVWAGEDRSKLANMATIRRI 180
Qy  181 DOLTNILAPMAGQIMTFSPIVGGFISGNWLSMCMVEYVLLWKYQKTPALAVKAGLK 240
Db  181 DOLTNILAPMAGQIMTFSPIVGGFISGNWLSMCMVEYVLLWKYQKTPALAVKAGLK 240
Qy  241 BEETELKQLNLHKOTPEKPLEGTHLMGVKDSNIHELEHQPTCASQMAEPFRFDGW 300
Db  241 BEETELKQLNLHKOTPEKPLEGTHLMGVKDSNIHELEHQPTCASQMAEPFRFDGW 300
Qy  301 SYYNQPVFLAGLAFYMTVLGFDCTTGAYTQGLSGSILSLMGASAITGIMGTAF 360
Db  301 SYYNQPVFLAGLAFYMTVLGFDCTTGAYTQGLSGSILSLMGASAITGIMGTAF 360
Qy  361 TWLRKCGLVRTGLISGLAQSLCLILCVISFMPGSPDLSPSPEDIRSFIOGESITP 420
Db  361 TWLRKCGLVRTGLISGLAQSLCLILCVISFMPGSPDLSPSPEDIRSFIOGESITP 420
Qy  421 TKIPITTEIYVMSNGSNANIVPETSPEVPTISVSLFAGVIAARIGLWSFDLTVTOLL 480
Db  421 TKIPITTEIYVMSNGSNANIVPETSPEVPTISVSLFAGVIAARIGLWSFDLTVTOLL 480
Qy  481 QENVIESERGIINGVQNSMNYLLDLHFIIMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
Db  481 QENVIESERGIINGVQNSMNYLLDLHFIIMVILAPNPEAFGLLVLSVSVFVAMGHIMYFR 540
Qy  541 FAQNTLGNKLPACGPDPAKEVKEKQNTSVV 571
Db  541 FAQNTLGNKLPACGPDPAKEVKEKQNTSVV 571

RESULT 4
Q9JH19 PRELIMINARY; PRT; 570 AA.
ID Q9JH19
AC Q9JH19
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE SLC11A3 iron transporter (ferroportin) (Solute carrier family 39
DE (iron-regulated transporter) (ferroportin), member 1).
GN FPN1 OR SLC39A1.
OS Mus musculus. (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=20347902; PubMed=10747949;
RA Abboud S., Haile D.J.;
RT "A Novel Mammalian Iron-regulated Protein Involved in Intracellular
RT Iron Metabolism";
RL J. Biol. Chem. 275:19906-19912(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20291023; PubMed=10828623;
RA Haile D.J.;
RT "Assignment of slc11a3 to mouse chromosome 1 band 1B and SLC11A3 to
RT human chromosome 2q32 by in situ hybridization.";
RL Cytogenet. Cell Genet. 88:328-329(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=20155474; PubMed=10693807;
RA Donovan A., Brownlie A., Zhou Y., Shepard J., Pratt S.J., Moynihan J.,
RA Paw B.H., Drejer A., Barut B., Zapata A., Law T.C., Brugnara C.,
RA Lux S.E. IV, Pinkus G.S., Pinkus J.L., Kingsley P.D., Palis J.,
RA Fleming M.D., Andrews N.C., Zon L.I.;
RT "Positional Cloning of Zebrafish Ferroportin1 Identifies a Conserved
RT Vertebrate Iron Exporter";
RL Nature 403:776-781(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Stomach;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Bokoff D., Bojunga N., Carninci P., de Bona M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nardone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:695-690(2001).
RN [5]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RT Submitted (FEB-2001) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AF215637; AAF80987.1; -
DR EMBL; AF226613; AAF36696.1; -
DR EMBL; AK008700; BAB25840.1; -
DR EMBL; BC003438; AAH03438.1; -
DR MGD; MGI:1315204; SLC39A1.
DR InterPro; IPR001993; Mitoch carrier.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
SQ SEQUENCE 570 AA; 62702 MW; 7125CC6171394A0A CRC64;

Query Match      89.4%; Score 2618.5; DB 11; Length 570;
Best Local Similarity 90.2%; Pred. No. 9.4e-193;
Matches 518; Conservative 18; Mismatches 31; Indels 7; Gaps 3;

Qy  1 MTRAGDHNRRGCCGSLADYLSAKFLYLGHSLSTWGRMWHFAVSFLVELYGNLSLL 60
Db  1 MTRAGDHNRRGCCGSLADYLSAKFLYLGHSLSTWGRMWHFAVSFLVELYGNLSLL 60
Qy  61 TAVYGLVAGSVLVLGAILGDWVDKNARLKVAQTSLVVQNVSVILCGIILMMVFLHKL 120
Db  61 TAVYGLVAGSVLVLGAILGDWVDKNARLKVAQTSLVVQNVSVILCGIILMMVFLHKL 120
Qy  121 LTMVHGWLTSYLIITIANIANLASTATAITIQRDWIVVWAGEDRSKLANMATIRRI 180

```

```
Db 121 LTMHGWLTVCYILITITIANIANLASTATAITIQDWMVWVAGENRSLADNNATIRRI 180
QY 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGWNLVSMCVFVLLWKVYQKTPALAVKAGLK 240
Db 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGWNLVSMCVFVLLWKVYQKTPALAVKAGLK 240
QY 241 EEBETELKQNLNKHDKTEPKLEGTHLMGVKDSNIHELEHEOEPTCASQMAEPFRTFRDGMV 300
Db 241 VESELKQLTSPKDETPKLEGTGTHLMGEKDSNIHELEHEOEPTCASQMAEPFRTFRDGMV 300
QY 301 SYYNQPVFLAGMGLAFLYMTVLGDFCITTCYAYTQGLSGSILSLMGASAITGIMGTVA 360
Db 301 SYYNQPVFLAGMGLAFLYMTVLGDFCITTCYAYTQGLSGSILSLMGASAITGIMGTVA 360
QY 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVFMVPGSPDLVSVPFEDIRSRFTQGSITP 420
Db 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVFMVPGSPDLVSVPFEDIRSRFTQGSITP 420
QY 421 -TKIPE--ITTEIYMSNGSNSANIVPETSPEVPIISVLSLAFAGVIAARIGLSFOLT 477
Db 421 TTIKIPETVFTTEHMSNMNSN---VHEMSTKPIPIVSVLSLAFAGVIAARIGLSFOLT 476
QY 478 QLLQENVIESERGIINGVQNSMNYLLDLHFIWVILAPNPEAFGLLVLISVSFVAMGHM 537
Db 477 QLLQENVIESERGIINGVQNSMNYLLDLHFIWVILAPNPEAFGLLVLISVSFVAMGHM 536
QY 538 YFRPAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
Db 537 YFRPAQNTLGNQIFVCGPDEKEVTDENQNTSVV 570
```

RESULT 5

```
Q9JKP4 ID Q9JKP4 PRELIMINARY; PRT; 570 AA.
AC Q9JKP4;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Iron-regulated transporter IREG1.
GN SLC39A1 OR IREG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20337919; PubMed=10882071;
RA McKie A.T., Marciani P., Rolfe A., Brennan K., Wehr K., Barrow D.,
RA Miret S., Bomford A., Peters T.J., Farzaneh F., Hediger M.A.,
RA Hentze M.W., Simpson R.J.;
RT "A novel duodenal iron-regulated transporter, IREG1, implicated in the
RT basolateral transfer of iron to the circulation.";
RL Mol. Cell 5:299-309(2000).
DR EMBL; AF231120; AAF4329.1; -.
DR MGD; MGI:1315204; SLC39A1.
DR InterPro; IPR001993; Mitoch_carrier.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
SQ SEQUENCE 570 AA; 62656 MW; C648A6C61C51EDCD CRC64;
```

Query Match 89.2%; Score 2612.5; DB 11; Length 570;
Best Local Similarity 90.1%; Pred. No. 2.7e-192;
Matches 517; Conservative 18; Mismatches 32; Indels 7; Gaps 3;

```
QY 1 MTRAGDHNRRGCCGLADYLTSAKFLLYLGHSLTWGDRMWHFAVSFVLVELYGNLSLL 60
Db 1 MTKARDQTHQEGCCGSLANLYTSAKFLLYLGHSLTWGDRMWHFAVSFVLVELYGNLSLL 60
QY 61 TAVYGLVWAGSVLVGLNIGDWVKNARKLVAQTSLVQVQNSVILCGIILMMVFLHKHEL 120
Db 61 TAVYGLVWAGSVLVGLNIGDWVKNARKLVAQTSLVQVQNSVILCGIILMMVFLHKHEL 120
QY 121 LTMHGWLTVCYILITITIANIANLASTATAITIQDWMVWVAGENRSLADNNATIRRI 180
Db 121 LTMHGWLTVCYILITITIANIANLASTATAITIQDWMVWVAGENRSLADNNATIRRI 180
QY 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGWNLVSMCVFVLLWKVYQKTPALAVKAGLK 240
```

```
Db 121 LTMHGWLTVCYILITITIANIANLASTATAITIQDWMVWVAGENRSLADNNATIRRI 180
QY 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGWNLVSMCVFVLLWKVYQKTPALAVKAGLK 240
Db 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGWNLVSMCVFVLLWKVYQKTPALAVKAGLK 240
QY 241 EEBETELKQNLNKHDKTEPKLEGTHLMGVKDSNIHELEHEOEPTCASQMAEPFRTFRDGMV 300
Db 241 VESELKQLTSPKDETPKLEGTGTHLMGEKDSNIHELEHEOEPTCASQMAEPFRTFRDGMV 300
QY 301 SYYNQPVFLAGMGLAFLYMTVLGDFCITTCYAYTQGLSGSILSLMGASAITGIMGTVA 360
Db 301 SYYNQPVFLAGMGLAFLYMTVLGDFCITTCYAYTQGLSGSILSLMGASAITGIMGTVA 360
QY 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVFMVPGSPDLVSVPFEDIRSRFTQGSITP 420
Db 361 TWLRKCGLVRTGLISGLAQLSCLILCVISVFMVPGSPDLVSVPFEDIRSRFTQGSITP 420
QY 421 -TKIPE--ITTEIYMSNGSNSANIVPETSPEVPIISVLSLAFAGVIAARIGLSFOLT 477
Db 421 TTIKIPETVFTTEHMSNMNSN---VHEMSTKPIPIVSVLSLAFAGVIAARIGLSFOLT 476
QY 478 QLLQENVIESERGIINGVQNSMNYLLDLHFIWVILAPNPEAFGLLVLISVSFVAMGHM 537
Db 477 QLLQENVIESERGIINGVQNSMNYLLDLHFIWVILAPNPEAFGLLVLISVSFVAMGHM 536
QY 538 YFRPAQNTLGNKLFACGPDPAKEVRKENQANTSVV 571
Db 537 YFRPAQNTLGNQIFVCGPDEKEVTDENQNTSVV 570
```

RESULT 6

```
Q923U9 ID Q923U9 PRELIMINARY; PRT; 570 AA.
AC Q923U9;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Ferroportin 1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Yeh K.-Y., Yeh M., Glass J.;
RT "Iron induces ferroportin 1 (FPT1) clustering and redistribution in
RT rat intestinal epithelial cells.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF394785; AAK77858.2; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
SQ SEQUENCE 570 AA; 62586 MW; 766786F22F054E94 CRC64;
```

Query Match 88.8%; Score 2600.5; DB 11; Length 570;
Best Local Similarity 89.9%; Pred. No. 2.3e-191;
Matches 514; Conservative 19; Mismatches 36; Indels 3; Gaps 2;

```
QY 1 MTRAGDHNRRGCCGLADYLTSAKFLLYLGHSLTWGDRMWHFAVSFVLVELYGNLSLL 60
Db 1 MTKSRDQTHQEGCCGSLANLYTSAKFLLYLGHSLTWGDRMWHFAVSFVLVELYGNLSLL 60
QY 61 TAVYGLVWAGSVLVGLNIGDWVKNARKLVAQTSLVQVQNSVILCGIILMMVFLHKHEL 120
Db 61 TAVYGLVWAGSVLVGLNIGDWVKNARKLVAQTSLVQVQNSVILCGIILMMVFLHKHEL 120
QY 121 LTMHGWLTVCYILITITIANIANLASTATAITIQDWMVWVAGENRSLADNNATIRRI 180
Db 121 LTMHGWLTVCYILITITIANIANLASTATAITIQDWMVWVAGENRSLADNNATIRRI 180
QY 181 DQLTNILAPMAVGQIMTFGSPVIGCGFISGWNLVSMCVFVLLWKVYQKTPALAVKAGLK 240
```

Db 181 DOLTNILAPMAVGQIMTGPVIGCGFISGNLVMSCVEYFLWKVYQKTPALAVKAALK 240
Qy 241 EETELKQNLNKHDKTEPKPLEGTHLMGVKDSNIHELEHQEPTCASQMAEPFRTPDCGW 300
Db 241 VEESELKQLTSPKDETEPKPLEGTHLMGVKDSNIRELECEQPTCASQIAEPFRFDGW 300
Qy 301 SYNOPVFLAGMLAFLYMTVLGFCITTGAYTQGLSGSILSLMGLASAITGIMGTVA 360
Db 301 SYNOPVFLAGMLAFLYMTVLGFCITTGAYTQGLSGSILSLMGLASAITGIMGTVA 360
Qy 361 TWLRKGLVRLTGLSGLAQLSCLILCVISVPMPSPLDLSVPEDIRSRFIQGESI-T 419
Db 361 TWLRKGLVRLTGLSGLAQLSCLILCVISVPMPSPLDLSVPEDIRSRFIHEAVSS 420
Qy 420 PTKIPEITTEIYMSNGSNSANIIVPETSPIISVLLFAGVIAARIGLWSFDLTVTQL 479
Db 421 TTKIPE--TEMLMSNVNVTVMHSTKSPFIISVLLFAGVIAARIGLWSFDLTVTQL 478
Qy 480 LOENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPAFGLLVLISVSVFVAMGHIMYF 539
Db 479 LOENVIESERGIINGVQNSMNYLLDLHLFIMVILAPNPAFGLLVLISVSVFVAMGHIMYF 538
Qy 540 RFAQNTLGNKLPACGPDADKEVRKENQANTSVV 571
Db 539 RFAQNTLGNKLPACGPDADKEVTDQNTSVV 570

RESULT 7
Q8JFWO PRELIMINARY; PRT; 562 AA.
ID Q8JFWO
AC Q8JFWO
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE SI:dz211013.7 (Solute carrier family 39 (Iron-regulated transporter), member 1).
DE member 1).
GN SLC39A1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN SEQUENCE FROM N.A.
RA Babbage A.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL591593; CAD3474.1; -.
SQ SEQUENCE 562 AA; 61761 MW; 35CB30971553718F CRC64;

Query Match 67.4%; Score 1975.5; DB 13; Length 562;
Best Local Similarity 71.0%; Pred. No. 2.3e-143;
Matches 394; Conservative 64; Mismatches 72; Indels 25; Gaps 10;
Qy 14 CGSLADYLTSAKFLYLGHSLSTWGRMWHFVAVSVFLVELYGNLSLLTAVYGLVAVGSVL 73
Db 11 CERPREFFKSAKFLIYVGHSLSTWGRMWHFVAVSVFLVELYGNLSLLTAVYGLVAVGSVL 70
Qy 74 VLGAIGDWVDKRNARKVAQTSLVVQNSVILCGIILMMVFLHKHLLTMTHTGWLTCY 133
Db 71 LLGAIGDWVDKRNARKVAQTSLVVQNSVILCGIILMMVFLHKHLLTMTHTGWLTCY 130
Qy 134 ILITIANIANLASTATAITORDWIVVAGDRSKLANNNATIRRIDLTNIIAPMAVG 193
Db 131 IWISIANIANLASTAMSIITORDWIVVAGDRSKLANNNATIRRIDLTNIIAPMAVG 190
Qy 194 QIMTFGSPVIGCGFISGNLVMSCVEYVLLWKVYQKTPALAVKAGLKE--EETELKQNLN 252
Db 191 QIMAFGSHFICGCGFISGNLVMSCLEFLYLLWKVYQKTPALAVKAGKQSDQDLKHLNIQ 250
Qy 253 KD---TEPKPLEGTHLMGVKDSNIHELEHQEPTCASQMAEPFRTPDCGWVSYNQPVFL 309
Db 251 KEIGNTE-SPVEASQLM-----TESSEPKKDTGCCQMAEPFRTPDCGWVSYNQVIF 303

Qy 310 AGMLAFLYMTVLGFCITTGAYTQGLSGSILSLMGLASAITGIMGTVAFTWLRKCGL 369
Db 304 AGMLAFLYMTVLGFCITTGAYTQGLSGSILSLMGLASAVSGICGTVAFTWLRKCGL 363
Qy 370 VRTGLISGLAQLSCLILCVISVPMPSPLDLSVPEDIRSRFIQGESI---IPTTKIPE 425
Db 364 IRTGFIAGVTQLSCLILCVASVAPGSPFDLSVPFEEV-LRHLFGDSGSLRESPTFIP- 421
Qy 426 ITTIYMSNGSNSANI-VPETSPIISVLLFAGVIAARIGLWSFDLTVTQLQSNV 484
Db 422 -TTTPPI-----QANVTVEEAPPVESYMSVGLLPAGVIAARVGLWSFDLTVTQLQSNV 475
Qy 485 IESERGIINGVQNSMNYLLDLHLFIMVILAPNPAFGLLVLISVSVFVAMGHIMYFRAQN 544
Db 476 IESERGIINGVQNSMNYLLDLHLFIMVILAPNPAFGLLVLISVSVFVAMGHIMYFRAQN 535
Qy 545 TLGNKLPFA-CGPDAD 558
Db 536 SLGSRLEFLFCSPQK 550

RESULT 8

Q9I9R3 PRELIMINARY; PRT; 562 AA.
ID Q9I9R3
AC Q9I9R3
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Ferroportin1.
GN SLC39A1 OR FPN1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN SEQUENCE FROM N.A.
RA TISSUE=Kidney;
RC MEDLINE=20155474; PubMed=10693807;
RA Donovan A., Brownlie A., Zhou Y., Shepard J., Pratt S.J., Moynihan J.,
RA Paw B.H., Drejer A., Barut B., Zapata A., Law T.C., Brugnara C.,
RA Lux S.E. IV, Pinkus G.S., Pinkus J.L., Kingsley P.D., Palis J.,
RA Fleming M.D., Andrews N.C., Zon L.I.;
RT "Positional Cloning of Zebrafish Ferroportin1 Identifies a Conserved
Vertebrate Iron Exporter";
RL Nature 403:776-781(2000).
DR EMBL; AF226612; AAF36695.1; -.
DR ZFIN; ZDB-GENE-000511-8; slc39a1.
SQ SEQUENCE 562 AA; 61760 MW; 38A6838E2107FBAF CRC64;

Query Match 67.3%; Score 1971.5; DB 13; Length 562;
Best Local Similarity 70.8%; Pred. No. 4.7e-143;
Matches 393; Conservative 65; Mismatches 72; Indels 25; Gaps 10;
Qy 14 CGSLADYLTSAKFLYLGHSLSTWGRMWHFVAVSVFLVELYGNLSLLTAVYGLVAVGSVL 73
Db 11 CERPREFFKSAKFLIYVGHSLSTWGRMWHFVAVSVFLVELYGNLSLLTAVYGLVAVGSVL 70
Qy 74 VLGAIGDWVDKRNARKVAQTSLVVQNSVILCGIILMMVFLHKHLLTMTHTGWLTCY 133
Db 71 LLGAIGDWVDKRNARKVAQTSLVVQNSVILCGIILMMVFLHKHLLTMTHTGWLTCY 130
Qy 134 ILITIANIANLASTATAITORDWIVVAGDRSKLANNNATIRRIDLTNIIAPMAVG 193
Db 131 IWISIANIANLASTAMSIITORDWIVVAGDRSKLANNNATIRRIDLTNIIAPMAVG 190
Qy 194 QIMTFGSPVIGCGFISGNLVMSCVEYVLLWKVYQKTPALAVKAGLKE--EETELKQNLN 252
Db 191 QIMAFGSHFICGCGFISGNLVMSCLEFLYLLWKVYQKTPALAVKAGKQSDQDLKHLNIQ 250
Qy 253 KD---TEPKPLEGTHLMGVKDSNIHELEHQEPTCASQMAEPFRTPDCGWVSYNQPVFL 309
Db 251 KEIGNTE-SPVEASQLM-----TESSEPKKDTGCCQMAEPFRTPDCGWVSYNQVIF 303


```
Db 251 KEIGNTE-SPVEASQLM-----TESSEPKKDKCCYQMAEPIRTFKDGMVAYNQSIFF 303
QY 310 AGMGLAFYMTVLGDFCITTYAYTQGLSGSILSLMGASAITGIMGTVAFTWLRKCGL 369
Db 304 AGMSLAFYMTVLGDFCITTYAYTQGLSGSVLSLLMGASAVSGICGTVAFTWIRKCGL 363
QY 370 VRTGLISGLAQLSCLILCVISVFMPSGLDLSVPPEDIRSRFTQGES-----ITPTKIPE 425
Db 364 IRTGFIAGVTQLSCLTLCVASVFAPGSPFDLSVSPFKEV-LRHLFGDSGLSRSPFTIP- 421
QY 426 ITTEIYMSNGSNANI-VPETSPESVPIISVSLFAGVIAARIGLMSFDTLTQTLLQENV 484
Db 422 -TTEPPI-----QANVTVEEAPPVESYMSVGLLFGACVIAARVGLMSFDTLTQTLLQENV 475
QY 485 IESERGIINGVQNSMNYLLDLHLHFMVILAPNPEAFGLLVLSVSVFVANGHIMYFRPAQN 544
Db 476 IESERGIINGVQNSMNYLLDLHLHFMVILAPNPEAFGLLVLSVSVFVANGHIMYFRPAYK 535
QY 545 TLGNKLFPA-CGPDAK 558
Db 536 SLGSRFLFCSPQK 550

RESULT 9
Q8AW23 ID Q8AW23 PRELIMINARY; PRT; 528 AA.
AC Q8AW23;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SI:d2182H3.5 (Solute carrier family 39 (Iron-regulated transporter), member 1) (Fragment).
GN SLC39A1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Kimberley A.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL672145; CAD58776.1; -.
FT NON TER 1
SQ SEQUENCE 528 AA; 57748 MW; 07ECE661DBE18851 CRC64;

Query Match 64.7%; Score 1894.5; DB 13; Length 528;
Best Local Similarity 71.8%; Pred. No. 3.5e-137;
Matches 381; Conservative 59; Mismatches 66; Indels 25; Gaps 10;

QY 38 GDRWHFAPSVFLVELYGNLSLLTAYGLVAGSVLVGLGAIIGDWVDKQARLKVAOTSLV 97
Db 1 GDRMNFAPSVFLVELYGNLSLLTAYGLVAGSVLVGLGAIIGDWVDKQARLKVAOTSLV 60
QY 98 VQNSVILCGIILMVLFLHKEHLLTWYHGVLTSCVILITIANIANLASTATAITQD 157
Db 61 VQNSAVILCGALLMVLVQFQKQSSWDGMLLTCTIWIISIANIANLASTATAITQD 120
QY 158 WIVVAGEDRSKLANMATIRRIDQLTNLAAPMAVGQIMTFGSPVIGCGFISGWNLSVMC 217
Db 121 WVVVAGDSDRSKLANMATVRIIDQLTNLAAPMAVGQIMAFGSHFICGCGFISGWNLSVMC 180
QY 218 VEYVLWKVYQKTPALAVKAGLKE-BETELKQNLKHD-----TEPKPLEGTHLMGVKDSNI 273
Db 181 LEYFLWKVYQKTPALAFKAGQSDSDQELKHLNIQKEIGNTE-SPVEASQLM-----T 233
QY 274 HELEHQEPTCASOMAEPPRTFDGMVSYNQVPELAGMGLAFYMTVLGDFCITTYAY 333
Db 234 ESSETKDKTCCYQMAEPIRTFKDGMVAYNQSIFFAGMSLAFYMTVLGDFCITTYAY 293
QY 334 TQGLSGSILSLMGASAITGIMGTVAFTWLRKCGLVRTGLISGLAQLSCLILCVISVFM 393
Db 294 TQGLSGSVLSLLMGASAVSGICGTVAFTWIRKCGLVRTGLISGLAQLSCLILCVASVFA 353
```

```
QY 394 PGSPDLDSVPPEDIRSRFTQGES-----ITPTKIPEITTEIYMSNGSNANI-VPETSP 448
Db 354 PGSPDLDSVPPEEV-LRHLFGDSGLSRSPFTIP--TTEPPI-----QANVTVEEAPP 405
QY 449 SVPIISVSLFAGVIAARIGLMSFDTLTQTLLQENVIESERGIINGVQNSMNYLLDLHLH 508
Db 406 VESYMSVGLLFGVIAARVGLMSFDTLTQTLLQENVIESERGIINGVQNSMNYLLDLHLH 465
QY 509 INVILAPNPEAFGLLVLSVSVFVANGHIMYFRPAQNTLGNKLFPA-CGPDAK 558
Db 466 INVILAPNPEAFGLLVLSVSVFVANGHIMYFRPAYKSLGSRFLFCSPQK 516

RESULT 10
Q8AW28 ID Q8AW28 PRELIMINARY; PRT; 436 AA.
AC Q8AW28;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SI:by113D7.1 (Solute carrier family 39 (Iron-regulated transporter), member 1) (Fragment).
GN SLC39A1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Sehra H.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL672118; CAD52128.1; -.
FT NON TER 1
SQ SEQUENCE 436 AA; 47761 MW; 3A1AE783CAB92FBC CRC64;
```

```
Query Match 51.7%; Score 1513.5; DB 13; Length 436;
Best Local Similarity 69.7%; Pred. No. 5.1e-108;
Matches 306; Conservative 51; Mismatches 57; Indels 25; Gaps 10;

QY 130 TSCYILITIANIANLASTATAITQDWMVIVVAGEDRSKLANMATIRRIDQLTNLAAP 189
Db 1 TTCYINVISIANIANLASTAMSIITQDWMVIVVAGDSDRSKLANMATVRIIDQLTNLAAP 60
QY 190 MAVGQIMTFGSPVIGCGFISGWNLSVMCVYVLLMKVYQKTPALAVKAGLKE-BETELKQ 248
Db 61 MLVGQIMAFGSHFICGCGFISGWNLSVMCLYELLWKVYQKTPALAFKAGQSDSDQELKH 120
QY 249 LNLKHD---TEPKPLEGTHLMGVKDSNIHELEHQEPTCASOMAEPPRTFDGMVSYNQ 305
Db 121 LNIQKEIGNTE-SPVEASQLM-----TESSETKDKTCCYQMAEPIRTFKDGMVAYNQ 173
QY 306 PVFLAGMGLAFYMTVLGDFCITTYAYTQGLSGSILSLMGASAITGIMGTVAFTWLR 365
Db 174 SIFFAGMSLAFYMTVLGDFCITTYAYTQGLSGSVLSLLMGASAVSGICGTVAFTWIRK 233
QY 366 KCGLVRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPEDIRSRFTQGES-----ITPT 421
Db 234 KCGLIRTGFIAGVTQLSCLTLCVASVFAPGSPFDLSVSPFEEV-LRHLFGDSGLSRSP 292
QY 422 KIPEITTEIYMSNGSNANI-VPETSPESVPIISVSLFAGVIAARIGLMSFDTLTQTLL 480
Db 293 FIP--TTEPPI-----QANVTVEEAPPVESYMSVGLLFGVIAARVGLMSFDTLTQTLL 345
QY 481 QNVIESERGIINGVQNSMNYLLDLHLHFMVILAPNPEAFGLLVLSVSVFVANGHIMYFR 540
Db 346 QNVIESERGIINGVQNSMNYLLDLHLHFMVILAPNPEAFGLLVLSVSVFVANGHIMYFR 405
QY 541 FAQNTLGNKLFPA-CGPDAK 558
Db 406 FAYKSLGSRFLFCSPQK 424
```



```

RESULT 11
Q921C9 PRELIMINARY; PRT; 405 AA.
AC Q921C9;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TReMBLrel. 10, Last annotation update)
DE Cell adhesion regulator.
GN CARL.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN RATTN010116;
RP RATTN010116;
RA Yang X.Z.;
SEQUENCE FROM N.A.
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U76714; AA000260.1; -.
SQ SEQUENCE 405 AA; 45145 MW; ABD5916C43AA4DA CRC64;

Query Match 50.5%; Score 1479.5; DB 11; Length 405;
Best Local Similarity 78.4%; Pred. No. 1.9e-105;
Matches 298; Conservative 15; Mismatches 36; Indels 31; Gaps 4;

Qy 1 MTRAGDHNRQRCGCCGLADYLSAKFLYLGLHSLSTWGDMMHFAVSVFLVELYXNSLL 60
Db 1 MTKSRDQTHQSGCCGLANLYLSAKFLYLGLHSLSTWGDMMHFAVSVFLVELYXNSLL 60

Qy 61 TAVYGLVAGSVLVGALIGDWDKNARKVAQTSVVQNVSVILCGIILMMVFLKHKL 120
Db 61 TAVYGLVAGSVLVGALIGDWDKNARKVAQTSVVQNVSVILCGIILMMVFLKHKL 120

Qy 121 LTMVHGWLTCYILITIANIANLSTATATITORDWVIVVAGEDRSKLANNNATIRI 180
Db 121 LNMTHGWLTCYILITIANIANLSTATATITORDWVIVVAGEDRSKLANNNATIRI 180

Qy 181 DQNLNAPMAGVQIMTFSPVIGCGFTSGWNLSVCMVEYVLLWKVQKTPALAVKAGLK 240
Db 181 DQNLNAPMAGVQIMTFSPVIGCGFTSGWNLSVCMVEYVLLWKVQKTPALAVKAGLK 240

Qy 241 BEETELKQNLHUKTEPKPLEGTHLMGVKDSNIHELEHQBPPTCASQWABPFRTRDGV 300
Db 241 VEESLQKLTSPKTEPKPLEGTHLMGVKDSNIHELEHQBPPTCASQWABPFRTRDGV 300

Qy 301 SYVNPQVFLAGWGLAF-----LYMTVLGFCITTG--YATQGLSGSILMGAS 349
Db 301 SYVNPQVFLAGWGLAF-----LYMTVLGFCITTG--YATQGLSGSILMGAS 349

Qy 350 AITGIMGTVAFTLRRKCGL 369
Db 352 -----NWNGNCGL 360

RESULT 12
Q8BME5 PRELIMINARY; PRT; 305 AA.
AC Q8BME5;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Solute carrier family 39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN MUSMUS10090;
RP MUSMUS10090;
RA Solute carrier family 39;
SEQUENCE FROM N.A.
RL STRAIN=C57BL/6J; TISSUE=Mesonephros;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK032732; BAC28001.1; -.
SQ SEQUENCE 305 AA; 33574 MW; EC9FE6264AF4D887 CRC64;

Query Match 46.4%; Score 1360.5; DB 11; Length 305;
Best Local Similarity 88.0%; Pred. No. 1.8e-96;
Matches 272; Conservative 11; Mismatches 19; Indels 7; Gaps 3;

Qy 266 MGKDSNIHELEHQBPPTCASQWABPFRTRDGVSVYVNPQVFLAGWGLAFLYMTVLGFD 325
Db 1 MGKDSNIHELEHQBPPTCASQWABPFRTRDGVSVYVNPQVFLAGWGLAFLYMTVLGFD 60

Qy 326 CITTGAYTQGLSGSILSMGASAITGIMGTVAFTWLRKCGLVRTGLISGLAQLSCLI 385
Db 61 CITTGAYTQGLSGSILSMGASAITGIMGTVAFTWLRKCGLVRTGLISGLAQLSCLI 120

Qy 386 LCVISVPMGSPDLDSVSPEDIRSRFIOGESITP-TKIPE--ITTEIYMSNGSNSANIV 442
Db 121 LCVISVPMGSPDLDSVSPEDIRSRFVNVEPVSPTTKIPETVFTTEHMSNMSN---V 176

Qy 443 PETSPEVPIISVSLIFAGVIAARIGLWSFDLTVTQLQENVIESERGIINGVONSNYL 502
Db 177 HEMSTRKPIPIVSVSLIFAGVIAARIGLWSFDLTVTQLQENVIESERGIINGVONSNYL 236

Qy 503 LDLLHFTIMVLAPNPAFAGLLVLSVSVFVAMGHIMYFRFAQNTLGNKLFACGPDKEVRK 562
Db 237 LDLLHFTIMVLAPNPAFAGLLVLSVSVFVAMGHIMYFRFAQNTLGNQIFVCGPDEKEVTD 296

Qy 563 ENQNTSVV 571
Db 297 ENQNTSVV 305

RESULT 13
Q9NUS1 PRELIMINARY; PRT; 167 AA.
AC Q9NUS1;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE Hypothetical protein FLJ11176.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sato H., Sugano S., Shiratori A., Sudo H.,
RA Wagaatsuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK002038; BAA92049.1; -.
DR InterPro; IPR001993; Mitoch_carrier.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
KW Hypothetical protein.
SQ SEQUENCE 167 AA; 18252 MW; 8D391F5356733660 CRC64;

Query Match 29.3%; Score 859; DB 4; Length 167;
Best Local Similarity 98.8%; Pred. No. 2.8e-58;
Matches 165; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 266 MGKDSNIHELEHQBPPTCASQWABPFRTRDGVSVYVNPQVFLAGWGLAFLYMTVLGFD 325
Db 1 MGKDSNIHELEHQBPPTCASQWABPFRTRDGVSVYVNPQVFLAGWGLAFLYMTVLGFD 60

Qy 326 CITTGAYTQGLSGSILSMGASAITGIMGTVAFTWLRKCGLVRTGLISGLAQLSCLI 385

```

```
Db 61 CTTGAYTQGLSGSILSLMGASAITGNGTVAFTWLRKCGLVRTGLISGLAQLSCLII 120
Qy 386 LCVISVFMPSGLDLSVSPEDIRSRPQIESITPTKIPITTEIYM 432
Db 121 LCVISVFMPSGLDLSVSPEDIRSRPQIESITPTKIPITTEIHW 167

RESULT 14
ID 080905 PRELIMINARY; PRT; 524 AA.
AC 080905;
DT 01-NOV-1998 (T-EMBLrel. 08, Created)
DT 01-NOV-1998 (T-EMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE At2g38460 protein.
GN At2G38460.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RC SEQUENCE FROM N.A.
RX MEDLINE=20083487; PubMed=106117197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Frazer C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004683; AAC28758.1; -.
SQ SEQUENCE 524 AA; 58204 MW; 781C6C0A697FADB1 CRC64;

Query Match 20.9%; Score 612.5; DB 10; Length 524;
Best Local Similarity 29.1%; Pred. No. 1.1e-38;
Matches 155; Conservative 95; Mismatches 179; Indels 103; Gaps 9;

Qy 28 LYLGHSLSTWGDMMHFAVSVELVLYGNSLLLTAVYGLVVGSLVVLGAILGDWDKNA 87
Db 41 LYVGFLRWASRTWFSVALYMIHLWPNLSLLAAIYGAIESGTALFGPIVGQWEGMD 100
Qy 88 RLKVAQTSLVQNVSVILCG---IILMMVFLHKHELLTMYHGWLTSCVILIITIANIAN 144
Db 101 YVKVLRLLWLFQNLSTYIAGGAVIKLLVSDLSKRNLPV-----FALLVLTNLAG 151
Qy 145 -----LASTATAITIQDWIVVVA-GEDRSKLANNNATIRIDQLTNILAPMAVGQIMTFG 199
Db 152 AIGVLTSLAGTILIERDWAVMSEGHPPAVLTQNSVIRGIDLSKLLSPVITGLIISFV 211
Qy 200 SPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALAVKAGLKEETELKQNLHKDTEPK 259
Db 212 SLKASAITFAAWATTATWVIEYWLFIISVGSVPAT-----RSNERRILSRTKQVEGRDAP 267
Qy 260 LEGTHLMGVKDSNIHELEHQEPTCASQMAEPRT-----FRDGNVSYNQF 306
Db 268 VSVISVPGTEEG-----YTGNP-----PSRTGILVILDRMSKSSFVGAWRIYFNQE 313
Qy 307 VFLAGGLAFLYMTVLGFCDCITTYAYTQGLSGSILSLMGASAITGNGTVAFTWLRK 366
Db 314 VLPFGVSLALLFTFTVLSFGTLMTATLQWEGIPTYYIIGIRGISAGVGLAATVLYPLMQSRI 373
Qy 367 CGLVRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPEDIRSRPQIESITPTKIPEI 426

Db 374 LSTLTGLWSFWSQSCLLVCVGSITWKKDK----- 404
Qy 427 TTEIYMSNGSNSANIVPETSPEVPIIISVSLFAGVIAARIGLWSFDLTVTQLLENVIE 486
Db 405 -----IASYMLMAGVAASRLGLMFMFDLAVIQMQDLVSE 438
Qy 487 SERGIINGVQNSMYLLDLLLHFIMVILAPNPEAFGLLVLSVSVFAMGHIMY 538
Db 439 SDRCVGVGVQNSLQALDLMAYLLGLIIVSNPKDFWILLISFSTVSLAGMLY 490

RESULT 15
QyLZC8 PRELIMINARY; PRT; 498 AA.
ID Q9LZC8
AC Q9LZC8;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (T-EMBLrel. 15, Last annotation update)
DE Transporter like protein.
GN F12B4370..
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RC SEQUENCE FROM N.A.
RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL162751; CAB83320.1; -.
SQ SEQUENCE 498 AA; 55504 MW; D8A4C5F61827998D CRC64;

Query Match 19.3%; Score 564; DB 10; Length 498;
Best Local Similarity 27.5%; Pred. No. 5.3e-35;
Matches 153; Conservative 102; Mismatches 175; Indels 126; Gaps 12;

Qy 28 LYLGHSLSTWGDMMHFAVSVELVLYGNSLLLTAVYGLVVGSLVVLGAILGDWDKNA 87
Db 37 LYLGFLLRWGARTWFSVALYMIHLWPNLSFLTAMYGVVSEGSATLFGPIVGQIMDGN 96
Qy 88 RLKVAQTSLVQNVSVILCG---IILMMVFLHKHELLTMYHGWLTSCVILIITIANIAN 144
Db 97 YVKVLRLLWLTQNLSTYIAGGAVALLVVPDLKSNQFPV-----FATLVLTNLSG 147
Qy 145 -----LASTATAITIQDWIVVVA-GEDRSKLANNNATIRIDQLTNILAPMAVGQIMTFG 199
Db 148 AIGVLTSLAGTILIERDWAVMSEGHSPAVLTQNSVIRGIDLSKLLSPVITGLIISFV 207
Qy 200 SPVIGCGFISGWNLSVCMVEYVLLWKVYQKTPALA---VKAGLKEEETELKQNL----- 250
Db 208 SLRASAITFAAWATTATWVIEYWLFIISVNGVPAIVQSDERRSLRSQSQAETDSASSFY 267
Qy 251 ---LHKDTEPKLEGTHLMGVKDSNIHELEHQEPTCASQMAEPRTFRDGNVSYNQPV 307
Db 268 VPLLHEESYRNTQS-----RSRILRILRISE-----SSFVSAMRNVLNQEI 310
Qy 308 FLAGGLAFLYMTVLGFCDCITTYAYTQGLSGSILSLMGASAITGNGTVAFTWLRK 367
Db 311 VLPFGVSLALLFTFTVLSFGTLMTATLQWEGIPTYYIIGIRGISAGVGLAATVLYPLMQSRI 370
Qy 368 GLVVRTGLISGLAQLSCLILCVISVFMPSGLDLSVSPEDIRSRPQIESITPTKIPEIT 427
Db 371 SPERTGVWSFSQK-----EKIAS----- 389
Qy 428 TTEIYMSNGSNSANIVPETSPEVPIIISVSLFAGVIAARIGLWSFDLTVTQLLENVIE 487
Db 390 ---YM-----LMAGVAASRLGLMFMFDLAVIQMQDLVPS 421
```

Search completed: November 5, 2003, 18:30:50
Job time : 62 secs

THIS PAGE BLANK (USPTO)